

The Japanese *A no B* Construction and Centering

Laurel Fais

NTT Communication Science Laboratories

NTT Corporation, Kyoto, Japan

fais@cslab.kecl.ntt.co.jp

Abstract

Applying a centering analysis to utterances containing complex noun phrases has always been problematic. This work examines the Japanese *A no B* construction, a complex noun phrase which exhibits a wide variety of semantic relations. A description is given of the types of relations found in examples of this structure taken from a corpus of Japanese email. Further, those examples which interact with a centering approach to the resolution of zero arguments in Japanese are investigated in detail to determine the relative prominence of the discourse elements *A* and *B* within that approach.

Introduction

Centering, one recent approach to modelling discourse, uses as basic elements the nominal arguments appearing in utterances, and its procedures depend upon well-defined roles played by those arguments. In structures in which roles are not well-defined, then, problems for the theory arise.

The Japanese *A no B* construction constitutes one of those structures. Used for possessive constructions, it also frames a wide variety of other semantic relationships, and its arguments, *A* and *B*, enter into centering processes. However, there is as yet no agreement as to how those arguments should be interpreted within the centering framework.

In this paper, I report on an examination of examples of the *A no B* construction found in a naturally occurring corpus of Japanese email. In the first section, I describe the corpus and the types of *no* phrases contained in it. Basic centering principles are outlined in the second section, along with a description of how the *A no B* examples in the corpus interact with those principles. In the final

section, I draw conclusions about how these constructions can most appropriately be incorporated into a centering framework.

1 *A no B* examples in the corpus

The data examined consist of a collection of 32 email messages exchanged among five employees of a Japanese company from June 5 to 16, 1995. The messages constitute a collective attempt to schedule a sports watching outing convenient for and interesting to all five in the group. Thus, the tone is usually casual. There is a total of 146 paragraphs, 228 sentences and 518 clauses in the corpus. Further details on the nature of the corpus can be found in Fais (2001).

The *A no B* constructions found in the corpus were classified according to the major types of semantic relations discussed in Shimazu, Naito, and Nomura (1987). They give a kind of case frame analysis of a wide range of *A no B* constructions. A brief description of some of the classifications, typical examples, and frequencies are given in Table 1 (next page)¹. In Cases 1 and 5, *A* modifies the head *B*, while in Cases 2 and 4, *B* modifies the head *A*. The interpretation of the whole phrase depends upon the semantics of *A* and *B* as well as on the case relationships they bear to one another.

2 Centering and the *A no B* construction

The centering framework as described in Walker, Joshi and Prince (1998a) is the foundation for the approach taken here. Each utterance in a discourse is considered to contain a set of discourse entities called

¹ This is not an exhaustive listing, merely a representative sample of the examples found in our corpus. For a complete description of the classification, see Shimazu, Naito and Nomura (1987).

Table 1. Classification of representative *A no B* examples in Japanese email corpus, number and percent.

Shimazu <i>et al.</i> 's classification	Example from corpus [Japanese] [transliteration]	[word-for-word rendering] [English translation]	#	%
Case 1 examples		Total in corpus:	76	54
B is predicate; A is subject	U-さん の 提案 U-san no teian	U-san <i>no</i> suggestion "U-san's suggestion"	23	16
B is predicate; A is object	ラクロス の 解説 lacrosse no kaisetsu	lacrosse <i>no</i> explanation "explanation of lacrosse"	14	10
B is predicate; A is time	17日 の 試合 17 nichi no shiai	17 th <i>no</i> game "game on the 17 th "	23	16
B is predicate; A is location	神宮球場 の ゲーム jingu kyujo no game	Jingu Stadium <i>no</i> game "game at Jingu Stadium"	13	9
Case 2: B is case role; A is head	ナイターの方 naita no hou	night game <i>no</i> direction "as for the night game"	33	23
Case 4 examples		Total in corpus:	3	2
A is predicate; B is subject	ヤクルト好きの 私 yakultozuki no watashi	Yakult liking <i>no</i> I "Yakult fan like me"	1	1
A is predicate; B is object	決行 の サッカー kekkoh no sakkah	holding <i>no</i> soccer "holding of a soccer game"	2	1
Case 5 examples		Total in corpus:	30	21
A restricts B; A is relational	私 の watashi no 誕生日会 tanjoubikai	I <i>no</i> birthday party "my birthday party"	12	8
A restricts B; A is situational	うしろ の おばさん ushiro no obasan	behind <i>no</i> woman "woman behind (me)"	11	8

forward-looking centers or Cfs. These entities are ranked in the Cf list for each utterance according to language-specific ranking principles. I follow the ordering principles for Japanese given in Walker, Iida, and Cote (1994):

(Grammatical OR \emptyset) topic > empathy > subject > object2 > object > others

This ranking, in effect, represents the salience of the discourse entities in an utterance, with topic-marked entities being most salient (for Japanese). Centering principles ensure that the referent for a pronominal or zero argument is the most salient entity in an utterance which also

appeared in the previous utterance. Notice that the entities involved in *A no B* constructions are not mentioned in the Cf ranking list given above, and thus their salience is undetermined in the theory.

This construction resembles the English possessive construction most closely, the arguments of which are also not included in the standard Cf rankings for English. A number of proposals have been made concerning possessor and possessed elements in centering accounts of English. According to the Complex Noun Phrase Constraint of Walker and Prince (1996), nominal elements are ranked from left to right as they appear in the complex noun phrase. On the other hand, Gordon *et al.*

(1999) claim, based on psycholinguistic experiments, that the possessed element (i.e., the rightmost element) should be ranked more highly. Tetreault (2001) tested these two hypotheses on two sets of texts and found that Walker and Prince's ordering works marginally better than Gordon *et al.*'s, results that were confirmed by Poesio and Nissim (2001).

A third approach is that of Di Eugenio (1998), who suggests that, in Italian, the possessor is ranked before the possessed if the possessed is inanimate, and immediately after it if the possessed is animate. Poesio and Nissim tested this approach as well (with the added provision that a pronominal possessor is more highly ranked) and found that this metric performed the best of all.

These studies can only be suggestive for the Japanese *A no B* construction. Yamura-Takei (personal communication) claims that Di Eugenio's animacy constraints accurately describe saliency for a subset of Japanese *A no B* genitive constructions. Nariyama (2000) has also argued for the crucial importance of animacy for resolving zero arguments in Japanese, but does not deal specifically with centering or with *A no B* constructions. Matsui (1999) discusses these Japanese constructions only to claim that both elements need to be accessible.

I examined the nine cases in our corpus in which entities in an *A no B* structure interact with centering processes by serving as referents for following zero arguments (Table 2 lists eight of them; the ninth is discussed below).

Table 2. Salient entities in *A no B* examples.

Type	#	Most salient
Case 1: A (time); B (head)	3	B (head)
Case 2: A (head); B (modifier)	2	A (head)
Case 1: A (subject); B (head)	1	A (subject)
Case 1: A (object); B (head)	1	A (object)
Case 4: A (head); B (object)	1	B (object)

In three examples of Case 1 in which *A* is a time entity (modifier) and *B* a nominal

form of a predicate (head), and in two examples of Case 2 in which *B* is the semantically thin *hou* (acting like an English preposition) and *A* is a head, the head entities served as referents for later zero arguments.

In three examples containing either subjects or objects acting as modifiers to the predicate (i.e., Case 1 (subject), Case 1 (object) and Case 4 (object)), the subject or object served as referent. Note that this is the *A* entity in Case 1 and the *B* in Case 4.

To account for the examples described up to this point, I argue that the head entity is the most salient entity in an *A no B* construction, unless the modifying entity fills an obligatory case role (such as subject and object). In that case, the modifying entity is the most salient.

This seems to be borne out by the remaining example. Consider [1].

- [1] ドームの日ハム戦 とか
 domu no nichihamusen toka
 Dome at Nichi Ham game or
マリンスタージアム (ちょっと遠いけど)
 marinsutajiamu \emptyset chotto toi kedo
 Marine Stadium \emptyset a bit far though
の千葉ロッテ戦 に行ってみる
 no chibarottesen ni ittemiru
 at Chiba Lotte game to try to go
 というのもありだと思います。
 toiuonomo aridato omoimasu
 another possibility think

"I think it would be another possibility to go to a Nichi Ham game at the Dome or to the Marine Stadium (though it is a bit far) for a Chiba Lotte game."

The *A no B* construction under consideration here is underlined; it is an example of a Case 1 (locative), and is split by a parenthetical. I have conjectured that in a Case 1 (locative), the most salient entity is the *B* entity, the head, since the locative case role is non-obligatory. If the phrase *though it is a bit far* had been positioned after the entire *no* phrase, the referent for the zero argument represented in English by *it* would have been a *Chiba Lotte game*, that is, the head of the phrase. The author, however, intended the *it* to

refer to *the Dome*, the optional case role argument. In order to make that reference successful, it was necessary for him to insert the zero argument immediately after *the Dome* and before *a Chiba Lotte game*.

All of the examples except one involve only inanimate elements; the exception is the Case 1 (subject) example, *U-san's work*. In this example, the predictions made by the animacy constraints discussed above and those made by the case role constraints are identical: *U-san* should be the more salient entity. This is borne out by the fact that *U-san* is, indeed, the referent for a following zero argument.

However, animacy constraints predict that in all cases in which *B* is inanimate, the *A* element is more salient. This conflicts with the example of Case 4 (object), *otazune no mise*, [recommendation *no mise*] "recommended shop." The *B* element, *shop*, is inanimate and so the *A* element *recommendation* should have more salience. However, it is *shop* that serves as the referent for a later zero argument. This follows from the argument above that an obligatory case element, i.e., the object *shop*, has more salience than the predicate.

3 Conclusion

I examined a corpus of Japanese email and categorized the types of *A no B* constructions found in order to assess empirically the validity of various claims about how these complex noun phrases interact with centering principles. My findings suggest that the nature of the case slot filled by an entity determines its salience: obligatory case entities are more salient than their predicates (heads), while non-obligatory case entities are less salient. There is a small amount of data to suggest that these considerations are more appropriate than animacy constraints suggested in the literature. However, much more data needs to be examined before a definitive conclusion can be reached. This is a promising area for future research.

References

Di Eugenio, Barbara. 1998. Centering in Italian. In Walker *et al.*, 1998b, 115-137.

- Fais, Laurel. 2001. Discourse issues in the translation of Japanese email. *Proceedings of PACLING 2001*, Kitakyushu, Japan, 93-102.
- Gordon, Peter C., Randall Hendrick, Kerry Ledoux, and Chin lung Yang. 1999. Processing of reference and the structure of language: An analysis of complex noun phrases. *Language and Cognitive Processes* 14(4):353-379.
- Matsui, Tomoko. 1999. On the role of context in relevance-based accessibility ranking of candidate referents. In P. Bouquet *et al.*, eds., *CONTEXT '99*, Springer Lecture Notes in Artificial Intelligence 1688: 228-241.
- Nariyama, Shigeko. 2000. Referent identification for ellipted arguments in Japanese. Ph.D. thesis, The University of Melbourne.
- Poesio, Massimo and Malvina Nissim. 2001. Saliency in possessive NPs: The effect of animacy and pronominalization. *AMLaP Conference Proceedings*, Saarbrücken.
- Shimazu, Akira, Shozo Naito, and Hirosato Nomura. 1987. Semantic structure analysis of Japanese noun phrases with adnominal particles. *25th Annual Meeting of the ACL, Proceedings*, Stanford, USA, 123-130.
- Tetreault, Joel R. 2001. A corpus-based evaluation of centering and pronoun resolution. *Computational Linguistics* 27(4).
- Walker, Marilyn A., Masayo Iida, and Sharon Cote. 1994. Japanese discourse and the process of centering. *Computational Linguistics* 20(2):193-231.
- Walker, Marilyn A. and Ellen F. Prince. 1996. A bilateral approach to givenness: A hearer-status algorithm and a centering algorithm. In Thorstein Fretheim and Jeanette K. Gundel, editors, *Reference and Referent Accessibility*, 89-112. John Benjamins Publishers, Amsterdam.
- Walker, Marilyn A., Aravind Joshi, and Ellen F. Prince. 1998a. Centering in naturally occurring discourse: An overview. In Walker *et al.*, 1998b, 1-28.
- Walker, Marilyn A., Aravind K. Joshi, and Ellen F. Prince, eds. 1998b. *Centering Theory in Discourse*. Clarendon Press, Oxford.