Anchoring Floating Quantifiers in Japanese-to-English Machine Translation

Francis Bond,[†] Daniela Kurz[‡] and Satoshi Shirai[†]
[†] NTT Communication Science Laboratories
2-4 Hikari-dai, Seika-cho, Soraku-gun, Kyoto, Japan, 619-0237
{bond,shirai}@cslab.kecl.ntt.co.jp
[‡] Department of Computational Linguistics, University of the Saarland
Postfach 1150, D-66041 Saarbrücken, Germany
kurz@coli.uni-sb.de

Abstract

In this paper we present an algorithm to anchor floating quantifiers in Japanese, a language in which quantificational nouns and numeral-classifier combinations can appear separated from the noun phrase they quantify. The algorithm differentiates degree and event modifiers from nouns that quantify noun phrases. It then finds a suitable anchor for such floating quantifiers. To do this, the algorithm considers the part of speech of the quantifier and the target, the semantic relation between them, the case marker of the antecedent and the meaning of the verb that governs the two constituents. The algorithm has been implemented and tested in a rule-based Japanese-to-English machine translation system, with an accuracy of 76% and a recall of 97%.

1 Introduction

One interesting phenomenon in Japanese is the fact that quantifiers can appear in two main positions, as pre-modifier in a noun phrase (1), or 'floating' as adjuncts to the verb phrase, typically in pre-verbal position (2).[1,2]

(1) *watashi-wa* ***3-ko-no*** *kēki-wo* *tabeta*
I-TOP 3-CL-ADN cake-ACC ate
I ate three cakes

(2) *watashi-wa* *kēki-wo* ***3-ko*** *tabeta*
I-TOP cake-ACC 3-CL ate
I ate three cakes

Quantifier 'float' of numeral-classifier combinations is widely discussed in the linguistic literature.[3] Much of the discussion focuses on identifying the conditions under which a quantifier can appear in the adjunct position. The explanations range from configurational (Inoue, 1983; Miyagawa, 1989) to discourse based (Downing, 1996; Alam, 1997), we shall discuss these further below. There has been almost no discussion of other floating quantifiers, such as quantificational nouns.

We call the process of identifying the noun phrase being quantified by a floating quantifier 'anchoring' the quantifier. The necessity of anchoring floating quantifiers for many natural language processing tasks is widely recognized (Asahioka et al., 1990; Bond et al., 1996), and is important not only for machine translation but for the interpretation of Japanese in general. However, although there are several NLP systems that incorporate some solution to the problem of floating quantifiers, to the best of our knowledge, no algorithm for anchoring floating quantifiers has been given. We propose such an algorithm in this paper. The algorithm uses information about case-marking, sentence structure, part-of-speech, noun and verb meaning. The algorithm has been implemented and tested within the Japanese-to-English machine translation system **ALT-J/E** (Ikehara et al., 1991).

The next section describes the phenomenon of quantifier float in more detail. We then propose our algorithm to identify and anchor floating quantifiers in Section 3. The results of implementing the algorithm in **ALT-J/E** are dis-

[1] Quantifiers are shown in **bold**, the noun phrases they quantify are underlined.

[2] This phenomenon exists in other languages, such as Korean. We will, however, restrict our discussion to Japanese in this paper.

[3] The name 'float' comes from early transformational accounts, where the quantifier was said to 'float' out of the noun phrase. Although this analysis has largely been abandoned, and we disagree with it, we shall continue with accepted practice and call a quantifier in the adjunct position a floating quantifier.

cussed in Section 4 and some remaining problems identified. The conclusion summarises the implementation of the algorithm and highlights some of its strengths.

2 Quantifier float in Japanese

First we will give a definition of quantifiers. Semantically, quantifiers are elements that serve to quantify, or enumerate, some target. The target can be an entity, in which case the number of objects is quantified, or an action, in which case the number of events (i.e. iterations of the action) are quantified. The quantification can be by a cardinal number, or by a more vague expression, like *several* or *many*.

In Japanese, quantifiers (**Q**) are mainly realised in two ways: numeral-classifier combinations (**XC**) and quantificational nouns (**N**). Note that these nouns are often treated as adverbs, as they typically function as adjuncts that modify verbs, a function prototypically carried out by adverbs. They can however head noun phrases, and take some case-markers, so we classify them as nouns.

Numeral classifiers form a closed class, although a large one. Japanese and Korean both have two or three hundred numeral classifiers (not counting units), although typically individual speakers use far less, between 30 and 80 (Downing, 1995, 346).

Syntactically, numeral classifiers are a subclass of nouns. The main property distinguishing them from prototypical nouns is that they cannot stand alone. Typically they postfix to numerals, forming a quantifier phrase, although they can also combine with the quantificational prefix *sū* "some" or the interrogative *nani* "what":

(3) *2-hiki* "2 animals" (Numeral)

(4) *sū-hiki* "some animals" (Quantifier)

(5) *nan-biki* "how many animals" (Interrogative)

Semantically, classifiers both classify and quantify the referent of the noun phrase they collocate with.

Quantificational nouns, such as *takusan* "much/many", *subete* "all" and *ichibu* "some", only quantify their targets, there is no classification involved.

Numeral classifier combinations appear in seven major patterns of use (following Asahioka et al. (1990)) as shown below (**T** refers to the quantified target noun phrase, **m** is a case-marker):

Type	Form	XC	N
pre-nominal	Q-no T-m	+	+
appositive	TQ-m	+	–
floating	T-m Q Q T-m	+	+
partitive	T-no Q-m	+	+
attributive	QT-m	+	–
anaphoric	T-m	+	–
predicative	T-wa Q-da	+	–

Table 1: Types of quantifier constructions

Noun quantifiers cannot appear in the appositive, attributive, anaphoric and predicative complement patterns.

In the pre-nominal construction the relation between the target noun phrase and quantifier is explicit. For numeral-classifier combinations the quantification can be of the object denoted by the noun phrase itself as in (6); or of a subpart of it as in (7) (see Bond and Paik (1997) for a fuller discussion). For nouns, only the object denoted by the noun itself can be quantified.

(6) ***3-tsū**-no tegami*
3-CL-ADN letter
3 letters

(7) ***3-mai**-no tegami*
3-CL-ADN letter
a 3 page letter

In the partitive construction the quantifier restricts a subset of a known amount: e.g., *tegami-no 3-tsū* "three of the letters". This is a very different construal to the pre-nominal construction. Only rational quantificational nouns can appear in the partitive construction.

The floating construction, on the other hand, has the same quantificational meaning as the pre-nominal. Two studies indicate that there are pragmatic differences (Downing, 1996; Kim, 1995). Pre-nominal constructions typically are used to introduce important referents, with non-existential predicates, while floating constructions typically introduce new number information. In addition floating constructions are used when the nominal has other modifiers, and are more common in spoken text.

We will restrict the following discussion to the difference between the pre-nominal and floating uses.

2.1 Restrictions on quantifier float

There have been many attempts to describe the situations under which the floating construction is possible, almost all of which only consider numeral-classifier constructions.

The earliest generative approaches suggested that the target in the floating construction must be either subject or object. Inoue (1983) pointed out that quasi-objects, noun phrases marked with the accusative case-marker but failing other tests for objecthood, could also be targets.

Miyagawa (1989) gives a comprehensive configurational explanation, where the target and quantifier must mutually c-command each other (that is, neither the target nor the quantifier dominates the other, and the first branching node that dominates either one, dominates the other). The restriction to nominative and accusative targets is explained by proposing a difference in structure. Verb arguments subcategorized for in the lexicon are noun phrases, where the case-marker is a clitic and thus can be c-commanded, whereas adjuncts are headed by their markers, to form post-positional phrases which are thus not available as targets.

The c-command relation is applied to both the noun phrases themselves and traces. Quantifiers can be scrambled (moved from their base position after their target) leaving a trace if the target is an *affected Theme* NP, and the target and quantifier are governed by the verb that assigns this thematic role. Thus quantifiers associated with affected themes can move within the sentence. Affected themes are things that are "changed, created, converted, extinguished, consumed, destroyed or gotten-rid of".

Miyagawa (1989, 57) proposes a syntactic test for affectiveness: affected themes can occure in the intransitive resultative construction *-te-aru*.

Alam (1997) looks at the problem from a different angle, and proposes that only quantifiers which are interpreted "distributively or as a quantified event" can float, as they take wide scope beyond the NP. A quantified noun phrase will also quantify the event if the noun phrase *measures-out* the event, where "direct internal arguments undergoing change in the event described by the verb measure out the event" a very similar description to that of **affected theme**. However, Jackendoff (1996) has shown that a wide variety of arguments can measure out processes, not just subjects and objects, but also the complements of prepositional phrases. Which case-roles measure out the process can be pragmatically determined as well as lexically stipulated, so it is not a simple matter to determine which arguments are relevent.

The excellent distributional analysis of Downing (1996) shows that actual cases of floating tend to be **absolutive**, that is quantifiers largely float from intransitive subjects (67%) or direct objects of transitive verbs (24%) rather than from transitive subjects (4%) or indirect objects (1%).

On the question of why quantifiers appear outside of the noun phrases they quantify, there have been two explanations: Discourse new information floats to the pre-verb focus position (Downing, 1996; Kim, 1995), quantifiers float from noun phrases that 'measure out' an event (Alam, 1997).

We speculate that there may be a performance based reason. Hawkins (1994) has shown that many phenomena claimed to be discourse related are in fact largely due to performance. However we have not yet compiled sufficient empirical evidence to show this conclusively.

3 An algorithm to identify and anchor floating quantifiers

The proposed algorithm is outlined in Figure 1. In our implementation it is appplied to each of one or more candidate outputs of a Japanese dependency parser as part of the semantic ranking.

3.1 Identify potential floating quantifiers

The first step is to identify potential floating quantifiers.

Every adjunct case element headed by a noun is checked. All numeral classifier combinations are potential candidates.

An adjunct must must meet two conditions to be considered a floating quantificational nouns, one semantic and one syntactic. The semantic criterion is that one of the noun's senses must be

```
For each unit sentence
   Identify potential floating
                quantifiers (QP)
    [Numeral-classifier
    or Quantificational Noun]
   Identify potential  anchors (NP)
    [nominative or accusative]
   Discard bad combinations
    [semantic anomalies,
     degree modifiers, event modifiers]
   Rank remaining combinations
      Prefer accusative
      Prefer anchor on the left
      Prefer closest
Anchor the best candidate pair(s)
```

Figure 1: Algorithm to anchor floating quantifiers

subsumed by `quanta, few/some, all-part`. The syntactic criterion is that the part of speech subcategory must be one of `degree` or `quantifier` adverbial.[4] We use the Goi-Taikei (Ikehara et al., 1997) to test for the senses and Miyazaki et al. (1995) for the syntactic classification.

3.2 Identify potential anchors

All noun phrases that matched a case-slot marked with *-ga* (nominative) or *-o* (accusative) are accepted as potential anchors. This is the traditional criterion given for potential anchors. Note even if the surface marker is different, for example when the case-marker is overwritten by a focus-marker such as *-wa* "topic", the 'canonical' case-marker will be found by our parser.

Noun phrases marked with *-ni* (dative), have been shown to be permissible candidates, but we do not allow them. Such sentences are, however, rare outside linguistics papers. We found no such candidates in the sentences we examined, and Downing (1996, 239) found only one in ninety six examples. When we tried allowing dative noun phrases, it significantly reduced the performance of our algorithm: every dative noun phrase selected was wrong. If we could determine which noun phrases measure-out the action, then they should also be considered as candidates, but we have no way to identify them at present.

3.3 Discard bad combinations

Some combinations of anchor and quantifier can be ruled out. We have identified three cases: semantically anomalous cases; sentences where the quantifier modifies the verb as a degree modifier; and sentences where the quantifier modifies the verb as a frequency modifier.

3.3.1 Semantically anomalous cases

Singular noun phrases In Japanese, pronouns and names are typically marked with a collectiviser (such as *-tachi*) if there are multiple referents (see e.g. Martin (1988, 143-154)). A pronoun or name not so marked characteristically has a singular interpretation. For names this can be overridden by a numeral-classifier combination (8), although it is rare, but not by an quantificational noun (9).

(8) *Matsuo-san-ga* ***3-nin*** *shabetta*
Matsuo-HON-NOM 3-CL spoke
3 Matsuos spoke

(9) *Matsuo-san-ga* ***takusan*** *shabetta*
Matsuo-HON-NOM many spoke
Matsuo spoke a lot

In all the texts we examined, we found no examples of names modified by floating numeral-classifier combinations. We therefore block all pronouns and names not modified by a collectiviser from serving as anchors to floating quantifiers.

In Japanese, there is not a clear division between pronouns and common nouns, particularly kin-terms such as *ojisan* "grandfather/old man". Pronouns can be modified in the same way as common nouns, and kin-terms are often used to refer to non kin. Pronouns modified by quantifiers need to be translated by more general terms as in (10).

(10) *kanojo-tachi-ga* ***3-nin*** *kita*
she-COL-NOM 3-CL came
? 3 she came
The 3 girls came

Classifier semantic restrictions For numeral classifiers, the selectional restrictions of the classifier can be used to disallow certain

[4] This part of speech category actually includes both true adverbs and adverb-like nouns.

combinations. For example, *-kai* "event" can only be used to modify event-nouns such as *shokuji* "meal" or *jishin* "earthquake". However, the semantics are very complicated, and there is a great deal of variation, as a classifier can select not just for the object denoted by its target but also a sub-part of it. In addition, classifiers can be used to select meanings figuratively, coercing a new interpretation of their head. Bond and Paik (1997) suggest a way of dealing with this in the generative lexical framework of Pustejovsky (1995) but it requires more information about the conceptual structure of noun phrases than is currently available.

For the time being, we use a simple table of forbidden combinations. For example *pointo* "point" will not be used to quantify nouns denoting `agent`, `place` or `abstract noun`.

3.3.2 Degree modification

Noun quantifiers can be used as degree modifiers as well as quantifying some referent. If the predicate is used to state a property of the potential anchor, then a noun quantifier will characteristically be a degree modifier.

We use the verbal semantic attributes given in the Goi-Taikei (Ikehara et al., 1997) to test for this relationship. Anchoring will be blocked either if the potential anchor is nominative and the verbal semantic attribute is one of `attribute transfer`, `existence`, `attribute` or `result` or if the anchor is accusative and the verbal semantic attribute is `physical/attribute transfer`.

Sentence (11) shows this constraint in action:

(11) *kodomo-ga* ***sukoshi*** *samui*
child-NOM a little cold

* A few children are cold

The child is a little cold

3.3.3 Event modification

The final case we need to consider is where the noun quantifier can quantify the event or the affected theme of the event, such as (12). In Japanese, either reading is possible when the quantifier is in pre-verbal position. Anchoring the quantifier is equivalent to choosing the theme reading.

(12) *kare-wa* *kēki-wo* ***takusan*** *tabeta*
he-TOP cake-NOM much ate

He ate cake a lot (event)

He ate a lot of cake (theme)

Examining our corpus showed the theme reading to be the default. Of course, if the event is modified elsewhere, for example by a temporal modifier, then different readings are possible. The system in which our implementation was tested lacks a system for event quantification, so we were not able to implement any constraint for this phenomenon. We therefore implemented the theme reading as our default. Note that, for stative verbs with permanent readings such as *shiru* "know", there is almost no difference between the two readings (13).

(13) *watashi-wa* *ratengo-wo* ***sukoshi***
I-TOP Latin-ACC a little
shitte-iru
know

I know a little Latin

I know Latin a little

3.4 Rank and select candidates

If there are more than two combinations, the following heuristics are used to choose which one or ones to choose.

Prefer accusative: A combination with an accusative anchor gets two points: This is to allow for the absolutive bias.

Prefer left anchor: If the anchor is to the left of the quantifier score it with one point: Quantifiers tend to float to the right of their anchors.

Prefer closest: Subtract one for each intervening quantifier: Closer targets are better.

Finally select the highest scoring combination and eliminate any combinations that include the chosen quantifier and anchor. If there is still a combination left (e.g. there were two quantifiers and two targets) then select it as well.

These heuristics rule out crossing combinations in the rare instances of two quantifiers and two candidates.

Floating	Anchored		Not anchored	
Quantifiers:	Good	Bad	Good	Bad
Nouns (**N**):	12	2	7	0
Num-Cls (**XC**):	16	7	11	1
Total:	28	9	18	1

Table 2: Test results

3.5 Anchoring

Once the best combinations are chosen, the quantifier can be anchored to its target. We consider the best way to represent this would be by showing the semantic relation in a separate level from the syntax, in a similar way to the architecture outlined by Jackendoff (1997).

Our implementation is in a machine translation system and we simply rewrite the sentence so that the floating quantifier becomes an prenominal modifier of its target, marked with the adnominal case-marker *-no*. The resulting modifier is labeled as 'anchored', to allow special processing during the transfer phase.

4 Results and Discussion

The algorithm was tested on a 3700 sentence machine translation test set of Japanese sentences with English translations, produced by a professional human translator. A description of the test set and its design is given in Ikehara et al. (1994).

Overall, 56 possible combinations were found and 37 anchored in 3700 sentences: Table 2. Of these, 9 were anchored that should not have been, and 1 was not anchored that should have been. The accuracy (correctly anchored/anchored) was 76% (28/37), and the recall (correctly anchored/should be anchored) was 97% (28/29).

The major source of errors was from parsing errors in the system as a whole. All of the badly anchored numeral-classifiers combinations were caused by this. In this case, the algorithm has not degraded the system performance, it would have been a bad result anyway.

There were three problems with the algorithm itself. In one case an anaphoric quantifier was mistaken as a floating quantifier, in another the verbal semantic attribute check for degree modification gave a bad result. Finally there was one case where the default blocking for semantic anomalies blocked a good combination.

Translation of floating quantifiers

Note that anchoring a floating quantifier is only the first step toward translating it. Special handling is sometimes needed to translate the anchored quantifiers.

For example, Japanese has some universal pronouns that can stand alone as full noun phrases (14) or act as floating quantifiers (15): e.g., *minna* "everyone", *zen'in* "all members". When they are anchored, the information about the denotation of the head carried by the pronoun is redundant, and should not be translated. A special rule is required for this.

(14) *minna-ga* *sorou*
everyone-NOM gather
All members gather.

(15) *membā-ga* ***minna*** *sorou*
members-NOM everyone gather
All the members gather.
*Everyone's members gather.

Further work

The proposed algorithm forms a solid base for extensions in various ways.

1. Combine it with a fuller system of event semantics.
2. Make the treatment of classifier-target semantics more detailed, so that inbuilt semantic restrictions can be used instead of a table of forbidden combinations.
3. Use the results of the algorithm to help choose between candidate parses and integrate it with the resolution of zero pronouns.
4. Test the algorithm on other languages, for example Korean.

5 Conclusion

We have presented an algorithm to anchor floating quantifiers in Japanese. The algorithm proceeds as follows. First identify potential floating quantifiers: either numeral classifier combinations or quantificational nouns. Then identify potential anchors: all accusative or nominative noun phrases. Inappropriate combinations are deleted, either because of a semantic mismatch between the target and quantifier, or because the quantifier is interpreted as a degree or event modifier. Finally, possible combinations are ranked, with the accusative candidate being the best choice, then the closest and leftmost. The algorithm is robust and uses the full power of currently available detailed semantic dictionaries.

Acknowledgments

The authors thank Tim Baldwin, Yukie Kuribayashi, Kyonghee Paik and the members of the NTT Machine Translation Research Group for their discussion and comments on this paper and earlier versions. The research was carried out while Daniela Kurz visited the NTT Communication Science Laboratories. Francis Bond is currently also enrolled part time as a doctoral candidate at the University of Queensland's Center for Language Teaching & Research.

References

Yukiko Sasaki Alam. 1997. Numeral classifiers as adverbs of quantification. In Ho-Min Sohn and John Haig, editors, *Japanese/Korean Linguistics*, volume 6, pages 381–397. CSLI.

Yoshimi Asahioka, Hideki Hirakawa, and Shinya Amano. 1990. Semantic classification and an analyzing system of Japanese numerical expressions. *IPSJ SIG Notes 90-NL-78*, 90(64):129–136, July. (in Japanese).

Francis Bond and Kyonghee Paik. 1997. Classifying correspondence in Japanese and Korean. In *3rd Pacific Association for Computational Linguistics Conference: PACLING-97*, pages 58–67. Meisei University, Tokyo, Japan.

Francis Bond, Kentaro Ogura, and Satoru Ikehara. 1996. Classifiers in Japanese-to-English machine translation. In *16th International Conference on Computational Linguistics: COLING-96*, pages 125–130, Copenhagen, August. (cmp-lg/9608014).

Pamela Downing and Michael Noonan, editors. 1995. *Word Order in Discourse*, volume 30 of *Typological Studies in Language*. John Benjamins.

Pamela Downing. 1995. The anaphoric use of classifiers in Japanese. In Downing and Noonan (Downing and Noonan, 1995), pages 345–375.

Pamela Downing. 1996. *Numeral Classifier Systems, the case of Japanese*. John Benjamins, Amsterdam.

John A. Hawkins. 1994. *A performance theory of order and constituency*, volume 73 of *Cambridge studies in linguistics*. Cambridge University Press, Cambridge.

Satoru Ikehara, Satoshi Shirai, Akio Yokoo, and Hiromi Nakaiwa. 1991. Toward an MT system without pre-editing – effects of new methods in **ALT-J/E**–. In *Third Machine Translation Summit: MT Summit III*, pages 101–106, Washington DC. (cmp-lg/9510008).

Satoru Ikehara, Satoshi Shirai, and Kentaro Ogura. 1994. Criteria for evaluating the linguistic quality of Japanese to English machine translations. *Journal of Japanese Society for Artificial Intelligence*, 9(4):569–579. (in Japanese).

Satoru Ikehara, Masahiro Miyazaki, Satoshi Shirai, Akio Yokoo, Hiromi Nakaiwa, Kentaro Ogura, Yoshifumi Ooyama, and Yoshihiko Hayashi. 1997. *Goi-Taikei — A Japanese Lexicon*. Iwanami Shoten, Tokyo. 5 volumes.

Kazuko Inoue, editor. 1983. *Nihongo-no Kihonkouzou (Basic Japanese Structure)*. Sanseido, Tokyo. (in Japanese).

Ray Jackendoff. 1996. The proper treatment of measuring out, telicity and perhaps even quantification in English. *Natural Language and Linguistic Theory*, 14:305–354.

Ray Jackendoff. 1997. *The Architecture of the Language Faculty*. MIT Press.

Alan Hyun-Oak Kim. 1995. Word order at the noun phrase level in Japanese: quantifier constructions and discourse functions. In Downing and Noonan (Downing and Noonan, 1995), pages 199–246.

Samuel E. Martin. 1988. *A Reference Grammar of Japanese*. Tuttle.

Shigeru Miyagawa. 1989. *Structure and Case Marking in Japanese*, volume 22 of *Syntax and Semantics*. Academic Press, Amsterdam.

Masahiro Miyazaki, Satoshi Shirai, and Satoru Ikehara. 1995. A Japanese syntactic category system based on the constructive process theory and its use. *Journal of Natural Language Processing*, 2(3):3–25, July. (in Japanese).

James Pustejovsky. 1995. *The Generative Lexicon*. MIT Press.

Zusammenfassung

In diesem Papier beschreiben wir einen Algorithmus zur Resolution von 'floating quantifiers' im Japanischen. Japanisch ist eine Sprache, in der quantifizierende Adverbien oder Kombinationen aus Numeral + Klassifikator von der Nominalphrase, für die sie quantifizieren getrennt werden können, d.h. sie müssen nicht in unmittelbarer linearer Abfolge stehen.

Der Algorithmus unterscheidet Grad- und Ereignismodifikatoren von Adverbialen, die für Nominalphrasen quantifizieren und resolviert den richtigen Antezedenten für jeden 'floating quantifier'. Zur Anbindung an die richtige Nominalphrase finden die folgenden Parameter Berücksichtigung: Wortart der Quantifikators und des Antezedenten, die semantische Relation zwischen diesen beiden, die Kasusmarkierungen des Antezedenten und die Semantik des Verbs, das sowohl den Quantifikator als auch dessen Antezedenten regiert.

Der Algorithmus wurde implementiert und in einem regel-basierten Japanisch/Englischem Übersetzungssystem evaluiert.

概要

本論文は日本語における遊離数量詞を付着させる方法を提案する。

副詞を、程度及び事象を表す数量名詞と名詞句を数量化する数量詞に分け、数量詞と名詞句を適切にリンクさせる。情報として考慮しているのは名詞と数量詞の品詞や意味関係、名詞句の格助詞、動詞の用言意味属性である。

この方法を日英機械翻訳システムに実装したところ、７６％の適合率、９７％の再現率が得られた。

요약

본 연구는 일본어의 유리수량어를 명사구에 부착시키는 방법에 대한 것이다.

부사를 정도및 사상 을 나타내는 수량어, 명사구를 수량화하는 수량어로 구분하여, 수량어와명사구를 적절히 링크시킨다. 이때 이용되는 정보로서는 명사와 수량어의 품사및 의미관계, 명사구의 각조사, 동사의 용언의미속성등이다.

이 방법을 일영기계번역 시스템에 이용한 결과, 76%의 적합률과 97%의 재현율을 얻었다.