Resolving Zero Anaphora in Japanese

Tadashi Nomoto and Yoshihiko Nitta
Advanced Research Laboratory, Hitachi Ltd.
2520, Hatoyama, Saitama 350-03, JAPAN
E-mail: {nomoto, nitta}@harl.hitachi.co.jp
Tel.+81-492-96-6111 Fax. +81-492-96-6006

Abstract
The paper presents a computational theory for resolving Japanese zero anaphora, based on the notion of discourse segment. We see that the discourse segment reduces the domain of antecedents for zero anaphora and thus leads to their efficient resolution.

Also we make crucial use of functional notions such as empathy hierarchy and minimal semantics thesis to resolve reference for zero anaphora [Kuno, 1987]. Our approach differs from the Centering analysis [Walker et al., 1990] in that the resolution works by matching one empathy hierarchy against another, which makes it possible to deal with discourses with no explicit topic and those with cataphora [Halliday and Hassan, 1990].

The theory is formalized through the definite clause grammar (DCG) formalism [Pereira and Warren, 1980, Gazdar and Mellish, 1989; Longacre, 1979].

Finally, we show that graphology i.e., quotation mark, spacing, has an important effect on the interpretation of zero anaphora in Japanese discourse.

1 Introduction
Over the past years, schemes like Focusing and Centering have dominated computational approaches to resolving anaphora [Sidner, 1983; Walker et al., 1990]. Their success derives from the utility they have in identifying salient discourse entities such as topic and thereby locating the antecedent for an anaphor. But they all suffer from the problem of directionality; they process the text (the list of sentences) from left to right, picking out focus along the way and see if a anaphor corefers with a focus already encountered. With the one-way processing, forward-looking pronouns (cataphora) are not possible to resolve. Since Japanese has great tolerance with forward reference, a proper theory of zero pronouns should meet the problem of directionality.

In what follows, we discuss some points about discourse segment and zero pronoun in Japanese. We begin by introducing the idea of discourse segment. Consider the pair:

(1) Taro-ga sara<1 i>-wo dasi, Hanako
    nom plate acc prepare-and
    -ga 02<i> ryori -wo morituketa.
    nom food acc arranged
    Taro prepared the plates, Hanako arranged food
    on them.

(2) Taro -ga sara<~>-wo dasi, Hanako<i> -wa
    nom plate acc prepare-and
    01<i> 02<k> ryori-wo morituketa.
    top
    Taro prepared the plates, Hanako arranged food.

Here, 02 represents a suppressed expression. It acts as an indirect object of the verb moritsuketa.\(^1\) 1 and 2 are morphologically identical except that 1 has ga (nominative marker) where 2 has wa (topic marker). But they differ widely in meaning: 1 implies that Hanako arranged food on the plates that Taro prepared, the reading 2 does not imply; in 2,

\(^1\)Here and throughout, we intend the term 01 to represent a zero pronoun for the subject, 02 for the indirect object, and 03 for the direct object.
Hanako could have arranged food on plates somebody other than Taro prepared. Now locating the difference will involve the notion of discourse segment. A discourse segment is defined as a set of sentences which appear in some region of text and which is delimited by a topic-particle wa. Thus 2 breaks up into two segments, a clause with Taro-ga and one with Hanako-wa;1, containing no wa-marked element, forms a segment by itself. Section 2.1 provides syntactic definitions for the discourse segment.

Another important feature of discourse segment is that of complying with the Minimal Semantics Thesis (MST) [Nomoto, 1992], a functional property that makes a segment cohere. The MST says, 'Assume as identical any pair of zero pronouns if it is part of some segment and does not occur as arguments for the segment's predicate.' Thus any pair of zero pronouns that fall into the domain of discourse segment are taken to be coreferential, unless they occur for the same predicate.2 Significantly, the MST is amenable to syntactic treatment.

In addition, we make use of the empathy hierarchy to choose between coreference relationships admitted by the MST. We specify a predicate for the empathy hierarchy and resolve zero anaphora by unifying one predicate's empathy hierarchy with another which occurs in the same segment. Since unification is a non-directional operation, we are able to treat forward as well as backward reference.

2 Theory
2.1 General
A discourse segment (DS) is a two-part structure consisting of head and body; a head is a nominal with a wa marking; a body is a set of sentences, which end with a period. Note that an adjunctive clause is not a sentence here, since it ends with connectives like -node because, -kara because/after, -to and-then, etc. Formally, we assume sentence has the following analyses, which are given in the DCG formalism.

\( \text{S} \rightarrow \text{C}^+, \text{N(pp:wa)} \).

\( \text{S} \rightarrow \text{C}^+, \text{N(pp:wa)}, \text{C}^+. \)

\( \text{S} \rightarrow \text{C}^+. \)

\( \text{C}^+ \) denotes one or more occurrences of clause, \( \text{C}^+ \) zero or more occurrences of clause, and \( \text{N(pp:wa)} \) denotes a wa-marked nominal; \( \text{pp:wa} \) specifies that the attribute pp (for postposition) has wa for the value.3 Let us define discourse segment by:

\[ \begin{align*}
\text{D} &\rightarrow \text{S}+. \\
\text{T} &\rightarrow \text{D}+(\text{head:}). \\
\text{S}(\text{head:X}) &\rightarrow \text{C}+, \text{N(morph:X,pp:wa)}. \\
\text{S}(\text{head:X}) &\rightarrow \text{C}^+, \text{N(morph:X,pp:wa)}, \text{C}+. \\
\text{S}(\text{head:}) &\rightarrow \text{C}+. \\
\end{align*} \]

2.2 Headed vs. Non-Headed Discourse
The discourse can be perfectly intelligible without an explicit topic or wa-nominal, which implies that a discourse segment may not be headed at all. It appears, however, that a discourse segment always comes out headed except when there is no head available in the text. In fact, a segment associates with a head nominal regardless of where it occurs in that segment.

\[ \begin{align*}
\text{Taro}_1\text{-wa} \text{ 01}_2\text{ 02}_j \text{ seki -wo uzutte} \\
\text{-ageta node, 01}_i\text{ 02}_j \text{ orei -wo} \\
\text{help because thank} \\
\text{iwareta. 01}_i\text{ chotto terekusa katta.} \\
\end{align*} \]

nominal is dislocated from the sentence and leaves a gap behind. In fact the analysis meets some difficulty in accounting for the wa-nominal having semantic control over a set of period-marked sentences. cf. [Mikami, 1960]. Ours, however, is free from the problem, as we see below.
Because Taro gave him/her a favor of giving a seat, he/she thanked Taro, who was slightly embarrassed.

Because Taro gave him/her a favor of giving a seat, he/she thanked Taro, who was slightly embarrassed.

Because Taro gave him/her a favor of giving a seat, he/she thanked Taro, who was slightly embarrassed.

8, 9 and 10 each constitute a discourse segment headed with Taro. A discourse can be acceptable without any head at all:

The speaker of 11, or watashi I would be the most likely antecedent for the elided subjects here; whoever gave the favor was thanked for the kindness. Let us say that a discourse is headed if each of its segments is headed, and non-headed, otherwise. Our assumption is that a discourse is either headed or non-headed, and not both (e.g. figure 1, figure 2). 5

Formally, this will be expressed through the value for the head attribute.

An empty-headed discourse expands into one segment; its head value will be inherited by each of the S-trees down below. Note that unification fails on

The Centering algorithm is not equipped to deal with cases like 9 and 10, where the backward-looking center Taro refers back to an item in the previous discourse.

It is interesting to note that a multiple-head discourse may reduce to a single-head discourse. This happens when discourse segments (DS) for a discourse, share an identical head, say, Taro and head-unifies with each other. In fact, such a reduction is linguistically possible and attested everywhere. Our guess is that a repeated use of the same wa-phrase may help the reader to keep track of a coreferent for zero anaphora.

As for this food, first pour in some water. Then add meat after 5 min.

We see that 14 constitutes a single discourse segment. According to the minimal semantics thesis, all of the zeros in the segment are interpreted as coreferential, which is consistent with the reading we have for the example. Here is a more complex discourse.
Taro went downtown to buy a clothing. Masako got it for her birthday present and she was very happy.

The first two zeros refer to Taro and the last two refer to Masako. But this is exactly what the MST predicts; 15 breaks up into two discourse segments, one that starts with Taro-wa and the other that starts with Masako-wa, so zeros for each segment become coreferential.

2.4 Empathy Hierarchy

It appears to be a fact about Japanese that the speaker of an utterance empathizes or identifies more with the subject than with the indirect object; and more with the indirect object than with the direct object [Kuno, 1987; Kuno and Kaburaki, 1977]. In fact, there are predicates in Japanese which are lexically specified to take an empathy-loaded argument; yaru give and kureru receive are two such. For yaru, the speaker empathizes with the subject, but with the indirect object, in the case of kureru.

The relevance of the speaker's empathy to the resolution problem is that an empathized entity becomes more salient than other elements in the discourse and thus more likely to act as the antecedent for an anaphor.

Taro gave Masako a favor in buying her a book. She still keeps it with care.

In 16, 01, subject of the second sentence, corefers with the indirect object Masako in the first sentence, which is assigned empathy by virtue of the verb kureta.

Formally, we define the empathy hierarchy as a function with three arguments:

\[
empathy(Z_1, Z_2, Z_3)
\]

With the definition at hand, we are able to formulate the lexical specification for kureru:

\[
V(\text{empathy}(\text{Arg}_2, \text{Arg}_1, \text{Arg}_3), \text{subject:Arg}_1, \text{object2:Arg}_2, \text{object:Arg}_3) \rightarrow \{\text{kureru}\}.
\]

yaru has the formulation like the following:

\[
V(\text{empathy}(\text{Arg}_1, \text{Arg}_2, \text{Arg}_3), \text{subject:Arg}_1, \text{object2:Arg}_2, \text{object:Arg}_3) \rightarrow \{\text{yaru}\}.
\]

Further, let us assume that variables in the empathy hierarchy represent zero pronouns. If a variable in the hierarchy is instantiated to some non-zero item, we will remove the variable from the hierarchy and move the items following by one position to the left; we might call it empathy shifting. Now consider the discourse:

(17) 01<ci> 02<j> hon -wo yatta -node, book acc favored because
01<ck> 02<ck> orei -wo iwarea. gratitude acc say cop

'Because he/she gave a book to him/her, he/she was thanked for it.'

(18) a empathy(01<ci>, 02<j>, _)

b empathy(01<ck>, 02<ck>, _)

18(1) corresponds to the empathy hierarchy for the first clause in 17; 18(b) corresponds to the hierarchy for the second clause. Unifying the two structures gives us the correct result: namely, 01<ci> = 01<ck>, and 02<j> = 02<ck>. Notice that zero items in the segment are all unified through the empathy hierarchy, which in effect realizes the Minimal Semantics Thesis. As it turns out, the MST reduces the number of semantically distinct zero pronouns for a discourse segment to at most three (figure 3). We conclude the section with a listing of the relevant DCG rules.

With the definition at hand, we are able to formulate the lexical specification for kureru:

\[
V(\text{empathy}(\text{Arg}_2, \text{Arg}_1, \text{Arg}_3), \text{subject:Arg}_1, \text{object2:Arg}_2, \text{object:Arg}_3) \rightarrow \{\text{kureru}\}.
\]

Further, let us assume that variables in the empathy hierarchy represent zero pronouns. If a variable in the hierarchy is instantiated to some non-zero item, we will remove the variable from the hierarchy and move the items following by one position to the left; we might call it empathy shifting. Now consider the discourse:

(17) 01<ci> 02<j> hon -wo yatta -node, book acc favored because
01<ck> 02<ck> orei -wo iwarea. gratitude acc say cop

'Because he/she gave a book to him/her, he/she was thanked for it.'

(18) a empathy(01<ci>, 02<j>, _)

b empathy(01<ck>, 02<ck>, _)

18(1) corresponds to the empathy hierarchy for the first clause in 17; 18(b) corresponds to the hierarchy for the second clause. Unifying the two structures gives us the correct result: namely, 01<ci> = 01<ck>, and 02<j> = 02<ck>. Notice that zero items in the segment are all unified through the empathy hierarchy, which in effect realizes the Minimal Semantics Thesis. As it turns out, the MST reduces the number of semantically distinct zero pronouns for a discourse segment to at most three (figure 3). We conclude the section with a listing of the relevant DCG rules.

With the definition at hand, we are able to formulate the lexical specification for kureru:

\[
V(\text{empathy}(\text{Arg}_2, \text{Arg}_1, \text{Arg}_3), \text{subject:Arg}_1, \text{object2:Arg}_2, \text{object:Arg}_3) \rightarrow \{\text{kureru}\}.
\]

Further, let us assume that variables in the empathy hierarchy represent zero pronouns. If a variable in the hierarchy is instantiated to some non-zero item, we will remove the variable from the hierarchy and move the items following by one position to the left; we might call it empathy shifting. Now consider the discourse:

(17) 01<ci> 02<j> hon -wo yatta -node, book acc favored because
01<ck> 02<ck> orei -wo iwarea. gratitude acc say cop

'Because he/she gave a book to him/her, he/she was thanked for it.'

(18) a empathy(01<ci>, 02<j>, _)

b empathy(01<ck>, 02<ck>, _)

18(1) corresponds to the empathy hierarchy for the first clause in 17; 18(b) corresponds to the hierarchy for the second clause. Unifying the two structures gives us the correct result: namely, 01<ci> = 01<ck>, and 02<j> = 02<ck>. Notice that zero items in the segment are all unified through the empathy hierarchy, which in effect realizes the Minimal Semantics Thesis. As it turns out, the MST reduces the number of semantically distinct zero pronouns for a discourse segment to at most three (figure 3). We conclude the section with a listing of the relevant DCG rules.
3 T-structure in Discourse

3.1 Embedding and Interleaving

In this section, we will illustrate some of the ways in which T-structure figures in Japanese discourse.

What we have below is a father talking about the health of his children.

Chichioya -wa 01 -te, father top laugh and
Taro -wa yoku kaze -wo hiku -n'desuyo. top often cold acc catch aux-polite
Kinou -mo 01 -kaze -wo hi'ite, 01 yesterday also cold acc catch
gakkou -wo yasu -n'da -n'desuyo. school acc take leave past aux-polite
Masako -wa 01 -gen'ldde, 01 -kaze top healthy cold
-wo hi'ita koto -ga arimas en. acc caught experiende nom occur aux-neg
01 -itsson sotode -masuyo," often outdoors play aux-polite
- to 01 -ittsa. comp said

"Taro often catches a cold. He got one yesterday again and didn’t go to school. Masako stays in a good health and has never been sick with flu. I often see her playing outdoors." Father said with a smile on his face.

Here are the facts: (a) zero anaphora occurring within the quotation (internal anaphora) are coreferential either with Taro or with Masako; (b) those occurring outside (external anaphora), however, all refer to chichioya; (c) chichioya has an anaphoric link which crosses over the entire quotation; (d) syntactically, the quoted portion functions as a complement for the verb -ittsa. It appears, moreover, that an internal anaphor associates itself with Taro in case it occurs in the segment headed with Taro, and with Masako in case it occurs in the segment headed with Masako. Then, since the quoted discourse consists of a set of discourse segments, it will be assigned to a T-structure. But the structure does not extend over the part -ittsa, which completes the discourse, for the 01 corefers with chichioya, and neither with Taro or Masako. This would give us an analysis like one in figure 4.

The following discourse shows that the T-structure can be discontinuous:

[a] "Masako -ga kinou sigoto - wo nom yesterday work acc
yasun'da -n'desuyo." [b] Hahaoya -wa -wa took leave aux-polite mother nom
01 -isu -ni suwara -to 01 hanashi chair on sit when tell
hazimeta [c] "Kaze -demo 01 -hi'ita -noka." began. cold acc caught question
[d] -to Chichioya -ga 03 -tasuneta. comp father nom asked

"Masako took a leave from the work yesterday.", Mother began to tell, as she sat on the chair. "Did she catch a cold?", asked Father.

01 corefers with Masako, so [c] forms a T-structure with [a]. But the two are separated by a narrative [b]. Similarly, the coreference between 03 and Hahaoya gives rise to a T-structure that spans [d] and [b], but there is an interruption by narrative [c] (figure 5).

Here and below we call a tree rooted at T a ‘T-structure’ and one rooted at D a ‘D-structure’.

3.2 Problem

There is a curious interaction between a paragraph-break and a T/D-structure. [Fujisawa et al., 1993], for instance, observes a strong tendency that Japanese zero anaphora are paragraph-bounded. The following is from Nihon Keizai Shinbun, a Japanese economics newspaper.

Kawamata Hideo -Mr. H. Kawamata Mr. H. Kawamata Sagami Tetsudo Sagami Railways
kaichou. [San-gatsu] Mik-ka gone juichi-ji chairman March 3rd day a.m. 11-hour
nijusan-pun, kokyuuzen no-tame 23-minute respiratory insufficiency due-to

\[N(\text{morph:} X, \text{pp:} wa), C(\text{empathy}(Z1, Z2, Z3))].\]
purposes; they same-index anaphora are distributed over the dis-
jor departure from the traditional analyses of zero
break has an effect of marking a segment for the
course, it seems rather obvious that a paragraph
rest appears as
[Zero-anaphora are made explicit here for expository
purposes; they are not part of the newspaper. The
rest appears as it does in the paper.] From the way
same-index anaphora are distributed over the dis-
course, it seems rather obvious that a paragraph
break has an effect of marking a segment for the
discourse.\textsuperscript{9} The present theory, however, fails to deal
with the situation like this; it simply assigns a single
DS structure to the discourse in question, giving a
wrong interpretation that zero anaphora present are
all coreferential. As it stands, nothing in the theory
provides for treating graphological marks such as a
paragraph break. Yet, it is unclear to us whether a
paragraph break is a signal for a T- or D-structure.

\textbf{4 Conclusion}

We have developed a computational theory for re-
solving zero anaphora in Japanese, drawing on the
results from previous works on Japanese discourse
[Kuno, 1987; Kuno and Kaburaki, 1977], etc. A ma-
jor departure from the traditional analyses of zero
anaphora lies in the reduction of the space of an-
tecedents for zero anaphora. This has been made
possible by adopting ideas like Discourse Segment,
Minimal Semantics Thesis and Empathy Hierarchy.
In particular, we have shown that the Minimal Se-
manics Thesis leads to reducing the number of an-
tecedents for a segment to at most three. Also shown
in the paper is that the resolution of zero anaphora
is part of parsing text, so no additional mechanism is
needed. Furthermore, the present theory compares
favorably with the previous schemes like Focusing
and Centering in that it is able to deal with forward-
and backward-looking anaphora by virtue of the way
unification operates on the empathy hierarchy.

Part of our discussion has touched on the effect of
graphology on the semantics of discourse. To date,
no significant research has been done on that area
of academic interests. The literature suggest that in
the written language, texts, i.e., cohesivediscourses,
are marked through a variety of linguistic and non-
linguistic means: non-alphanumeric characters (quo-
tation marks, brackets, parentheses), graphic devices
(indentation, tabulation, itemization), and so on
[Nunberg, 1990; Halliday and Hassan, 1990]. Thus
a discourse segment might qualify for the texthood
since it has the property that zero pronouns are res-
olved internally. Its indicator is, of course, the topic
particle wa. But for the T-structure, it is far from
clear whether it is anyway cohesive, and if it is, what
its indicators are. (Quotation mark and paragraph
break are possible candidates.)

Some of the technical as well as linguistic details
are yet to be worked out; we have not talked about
how the topic comes to be associated with one or
more zero pronouns in the segment. Considerations
on empathy may well influence the choice of pro-
nouns to be matched with.

\textbf{References}

[Fujisawa et al., 1993] Shinji Fujisawa, Shigeru Masu-
yama, and Shozo Naito. An Inspection on Ef-
flect of Discourse Contraints pertaining to Ellip-
sis Supplement in Japanese Sentences. In \emph{Kouen-
Ronbun-Shuu 3 (conference papers 3)}. Information
Processing Society of Japan, 1993. In Japanese.

[Gazdar and Mellish, 1989] Gerald Gazdar and Chris Mellish.
Natural Language Processing in Prolog. Addison-Wesley Publishing Co., New
York, 1989.

[Gunji, 1987] Takao Gunji. \emph{Japanese Phrase Struc-
ture Grammar}. D. Reidel, Dordrecht, 1987.

[Halliday and Hassan, 1990] M. A. K. Halliday and R.
Hassan. \emph{Cohesion in English}. Longman, New
York, 1990.

[Hobbs, to appear] Jerry R. Hobbs. On the Coher-
ence and Structure of Discourse. in \emph{The Structure
of Discourse}, Livia Polanyi, editor, Ablex Publish-
ing Co., to appear.

[Inoue, 1978] Kazuko Inoue. \emph{Nihongo-no Bunpo
Kisoku ( Grammatical Rules in Japanese )}. Taishukan, Tokyo, 1978. In Japanese.

[Kitagawa, 1982] C. Kitagawa. Topic construction in
Japanese. \emph{Lingua}, 57:175–214, 1982.
[Kuno and Kaburaki, 1977] Susumu Kuno and Etsuko Kaburaki. Empathy and Syntax. *Linguistic Inquiry*, 8:627–672, 1977.

[Kuno, 1987] Susumu Kuno. *Functional Syntax*. The University of Chicago Press, Chicago, 1987.

[Kuroda, 1965] S. Y. Kuroda. *Generative Semantic Studies in the Japanese Language*. Garland, New York, 1965.

[Longacre, 1979] R. E. Longacre. The paragraph as a grammatical unit. In Tamly Givón, editor, *Syntax and Semantics vol. 12*. Academic Press, 1979.

[Mikami, 1960] Akira Mikami. *Zou wa Hana ga Nagai (The elephant has a long trunk).* Kuroshio Shuppan, Tokyo, 1960.

[Nomoto, 1992] Tadashi Nomoto. Discourse and semantics of zero-pronominals. In *Proceedings of NLC workshop, Nagasaki*, 1992.

[Nunberg, 1990] Geoffrey Nunberg. *The Linguistics of Punctuation*, volume 18 of *CSLI Lecture notes*. CSLI, 1990.

[Pereira and Warren, 1980] Fernando C. N. Pereira and David H. D. Warren. Definite clause grammar for language analysis - a survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence*, 13:231–278, 1980.

[Sidner, 1983] Candance L. Sidner. Focusing in the comprehension of definite anaphora. In Brady and Berwick, editors, *Computational Model of Discourse*, pages 267–330. The MIT Press, Cambridge, 1983.

[Walker et al., 1990] M. Walker, M. Iida, and S. Cote. Centering in Japanese. In *Proceedings of COLING '90*, 1990.