Approaches to Japanese zero pronouns: Centering and relevance

Tomoko Matsui
International Christian University
Division of Languages
3-10-2 Osawa
Mitaka, Tokyo, JAPAN 181
matsui@icu.ac.jp

Abstract

Accessibility of candidate referents is a crucial factor in successful reference assignment and this intuition has been adequately accommodated and developed in centering theory. By contrast, another equally important factor involved in reference assignment, namely, accessibility of contextual assumptions, seems to have not been addressed fully as a part of a computational theory of reference. In the field of pragmatics, however, the role of context has always been one of the central issues and the recent approach to context selection based on the notion of relevance (Sperber & Wilson 1986/1995) seems to be currently the most promising. In this paper, I will review and assess the recent centering approach to the interpretation of Japanese zero pronouns (Walker et al. 1994) as a case study, and suggest that relevance theory can provide one way of complementing it. In order to incorporate a needed mechanism of contextual selection into a model of reference assignment, model builders must take into account the hearer's ability to anticipate certain cognitive effects that may be achieved by upcoming utterances. It is my hope that this paper will lead to a fruitful discussion of finding possible ways of using Sperber & Wilson's notion of relevance in practical applications.

1. Introduction

When a system has to assign a referent to a referring expression, it is almost always the case that there is more than one candidate referents and one has to resort to some way of eliminating the wrong candidates in order to choose the right one.

There are two main ways of doing this. First, candidate referents can be ranked on the basis of accessibility (Erku & Gundel 1987; Fretheim and Gundel 1996; Gernbacher and Hargreaves 1988, Grosz et al. 1995; Sidner 1983a, b; Walker et al. 1998). Alternatively, they can be checked against the accessibility of contextual assumptions (Fincher-Kiefer 1993; Kintsch 1988, Magliano et al. 1993; McKoon & Ratcliff 1992; Sanford & Garrod 1981; Sharkey & Sharkey 1987; Singer 1993, 1995).

In either case, once a candidate is singled out, the acceptability of the referent needs to be tested against some pragmatic criteria. Currently, we seem to have at least three distinct criteria available:

i. **Truth-based criterion** – i.e. whether the overall interpretation is likely to be factually plausible (Clark 1977; Clark & Haviland 1977; Erku & Gundel 1987; Sanford & Garrod 1981; Sidner 1983a);
ii. **Coherence-based criterion** – i.e. whether the overall interpretation is likely to be coherent (Asher & Lascaries 1993; Grosz et al. 1995, 1998; Hobbs 1979; Lascarides & Asher 1993; Sanders et al. 1992; Walker et al. 1994, 1998);
iii. **Relevance-based criterion** – i.e. whether the overall interpretation is likely to be optimally relevant (Matsui 1993, 1995, 1998; Wilson 1992; Wilson & Matsui 1998).

11

In this paper, I will focus on one version of coherence-based criterion, namely, centering theory. One of the goals of centering theory is to sort out the various mechanisms used to maintain discourse coherence, and the use of various referring expressions is regarded as one such mechanism. Among the various hypotheses put forward by centering theory, what concerns us most is the following: that 'each utterance [except the initial utterance] in a coherent discourse segment contains a single semantic entity - the backward-looking center [or Cb] - that provides a link to the previous utterance, and an ordered set of entities - the forward-looking centers [or Cf] - that offer potential links to the next utterance '(Gordon et al. 1993:311). There are two rules to provide constraints on choosing centers, which are shown in (1):

(1)Rule 1: If any element of Cf (Un) is
    realised by a pronoun in Un+1 then the
    Cb (Un+1) must be realised by a
    pronoun also.
    Rule 2: Sequences of continuation are
    preferred over sequences of retaining;
    and sequences of retaining are to be
    preferred over sequences of shifting.

The first rule states that the most highly ranked element of the forward-looking center of a previous utterance is the backward-looking center of the current utterance, and must be realised by a pronoun if any element of the Cf of the previous utterance is realised by a pronoun in the current utterance. The following example from Gordon et al. (1993), shown here in (2), illustrates this rule:

(2)[1]Susan gave Betsy a pet hamster.
    Cf={Susan, Betsy, hamster 1}
    [2]She reminded her such hamsters were
        quite shy.
    Cb=Susan; Cf={Susan, Betsy, hamsters}
    [3]She asked Betsy whether she liked
        the gift.
    Cb=Susan; Cf={Susan, Betsy,
    gift=hamster 1}
    [3']Susan asked her whether she liked
        the gift.
Here, the first utterance has no Cb because it is

the initial sentence of a discourse. Its Cf includes the referents of 'Susan' and 'Betsy' and the semantic interpretation of 'a hamster', ranked in that order. The second utterance has Susan as the Cb and a Cf with Susan as its most highly ranked element. The third utterance preserves the Cb and prominent Cf from the previous utterance, therefore it pronominalises the Cb successfully. By contrast, utterance [3'], in which Susan is realised by a name and Betsy is realised by a pronoun, leads to stylistic infelicity. According to Gordon et al., this is due to violation of Rule 1 mentioned above.

The ranking of forward-looking centers is generally based on the discourse salience of each candidate entity. According to Grosz et al. (1995), although an ultimate criteria for deciding the ranking has not been worked out yet, there are evidences to support the idea that grammatical role such as SUBJECT, OBJECT, etc., can affect the Cf ranking. Thus, I will simply assume here the following preference in ranking forward-looking center shown in (3), as suggested by Grosz et al. (ibid.):

(3)SUBJECT>OBJECT>OTHERS

The highest ranked member of the set of forward-looking centers is called the 'preferred center', or 'Cp'. As mentioned above, Cp is regarded as the most likely candidate for Cb in the following utterance.

Another important claim made by centering theory is that discourse segments are more coherent if they share the same Cb. On the basis of this idea, different degrees of coherence are proposed. For example, Walker et al. adopt the following 4 types of transition between discourse segments, each corresponding to different degree of coherence, using the notion of Cb and Cp; namely, 'continue', 'retain', 'smooth-shift' and 'rough-shift'. These are shown in (4). When two utterances, say Ui-1 and U share the same Cb, and the same entity is also the highest-ranked Cf, i.e. Cp, in Ui-1, the transition from Ui-1 to Ui is called 'continue'. When Ui-1 and Ui share the same Cb, but the

same entity is not the highest-ranked Cf in Ui, the transition is called 'retain'. When Ui-1 and Ui do not share the same Cb, there are two possibilities: if Cb in Ui is the same as Cp in Ui, the transition is 'smooth-shift'; if Cb in Ui is not the same as Cp in Ui, it is 'rough-shift'. It is claimed that when the hearer has to choose one from several possible interpretations, the one based on the most coherent transition should be chosen. The 4 transition states are ordered in the following way according to their preference:

(4)CONTINUE>RETAIN>SMOOTH-SHIFT>ROUGH-SHIFT

The second of the two rules in (1) is about this ordered preference of transition states.[1]

2. Walker et al. (1994) and Japanese zero pronoun resolution

Walker et al. (1994) propose the following ranking order of forward-looking centers to deal with Japanese:

(5)(Grammatical or Zero) TOPIC > EMPATHY > SUBJECT > OBJECT2 > OBJECT> OTHERS

As you can see, they add two new grammatical roles, namely, topic and empathy, to the list of factors affecting pronoun resolution. Since the notion of topic will become important in the discussion which follows, I will briefly describe the Japanese topic marker below. For empathy-loaded verbs, please refer to Kuno 1987 and Kuno & Kaburaki 1977.

In Japanese, both in written and spoken discourse, NPs which can be recovered from context are often omitted. The omitted NPs are often called 'zero pronouns'. It is widely agreed that missing NPs in Japanese behave like pronouns in other languages such as

English. Japanese zero pronouns should be distinguished from missing NPs in 'pro-drop' languages such as Italian, since in 'pro-drop' languages, information to recover missing NPs is morphologically encoded elsewhere in the sentence, e.g. in the form of verb inflection, whereas Japanese lacks such an overt encoding.

Japanese zero pronouns are used equally often as inter-sentential discourse anaphors and intra-sentential anaphors. Here, I will concentrate on their use as discourse anaphors, where the role of context and adequate pragmatic criterion is crucial. For those who are interested in intra-sentential use of zero pronouns, please refer to e.g. Hasegawa 1985 and Kameyama 1986.

In Japanese, there are several postpositional particles. The one we are interested in here is 'wa', which is often called a 'topic marker'. As the name suggests, 'wa' is typically used to construct a grammatical topic of a sentence, which is characterised as an entity whose existence is presupposed. The function of 'wa' might become clearer when it is compared with another particle 'ga', which marks a NP in the subject position, which typically conveys a new information. Compare (6a) and (6b):

(6) a. *John ga hana   o   katta.*
        SUB flowers ACC bought
        'John bought flowers.'
    b. *John wa hana   o   katta.*
        TOP flowers ACC bought
        'John bought flowers.'

(6a) and (6b) share the same propositional content. However, only (6a) is acceptable as an answer to the question such as 'Who bought flowers?' while only (6b) is acceptable as an answer to the question such as 'What did John buy?'. There are various suggestions about how to characterise functions of the two particles (see e.g. Shibatani 1990; Tanaka 1991), and although it is an interesting question on its own, it shouldn't concern us here. This is because Walker et al. are only interested in the surface form of 'NP+wa',

---

[1] *There were only three transition states in the original formulation by Grosz et al. (1983, 1995), namely, CONTINUE, RETAIN and SHIFT. The distinction between SMOOTH-SHIFT and ROUGH-SHIFT was first proposed in Brennan (1987).*

which is automatically given the highest accessibility ranking in their framework.

Walker et al. claim that a topic NP is more likely to be realised as a zero pronoun in the subsequent discourse than any other NPs due to its highest degree of accessibility. Moreover, in Walker et al.'s framework, topic NP is given two further advantages: they suggest (a) that a NP marked by 'wa' becomes the backward-looking center even at the onset of the discourse; and (b) that once topic NP is realised as the backward-looking center ,as long as it continues to be realised as a zero pronoun in subsequent discourse, it could continue to be the backward-looking center. The second advantage given to the topic NP is called 'zero topic assignment', which is defined as in (7):

(7)Zero Topic Assignment (ZTA)
(optional)
When a zero in Ui+1 represents an entity that was Cb (Ui), and when no other CONTINUE transition is available, that zero may be interpreted as the ZERO TOPIC of Ui+1

Walker et al. demonstrate that their framework, including an optional rule of zero topic assignment, can successfully explain the preference in the interpretation of the last sentence in (8):

(8)[1] *Hanako wa siken o oete,*
   TOP exam ACC finish
   kyooshitu ni modorimashita.
   classroom to returned
   'Hanako returned to the classroom, finishing her exams '
   Cb=Hanako; Cf={Hanako, exam}
[2] 0 *hon o   rokka ni shimaimashita.*
   book ACC locker in took-away
   '(She) put her books in the locker '
   Cb=Hanako; Cf={Hanako, book}
   [CONTINUE]
[3] *itsumo no yooni Michiko ga*
   always like   SUB
   0 *mondai no tokikata   o*
   questions solve-way ACC

setumeishimashita.
   explained
   'Michiko, as usual, explained (to her) how to answer questions. '
   Cb=Hanako;
   Cf1={Hanako, Michiko ...} ZTA
   [CONTINUE]
   Cf2={Michiko, Hanako, ...}
   [RETAIN]
[4] 0 0 *ohiru ni sasoimashita*
          lunch to invited
   '(She) invited (her) to lunch '
   Cb1= Hanako;
   Cf1={Hanako, lunch, Michiko}
   [CONTINUE]from [3]-Cf1
   Cb2=Michiko
   Cf2={Michiko, lunch, Hanako}
   [S-SHIFT]from [3]-Cf2

According to the questionnaire carried out by Walker et al., the preferred interpretation of [4] is that Hanako invited Michiko to lunch. As you can see, in fact, there are two possible ways of ranking forward-looking center in [3], and two possible ways of deciding both the backward-looking center and the ranking of forward-looking center in [4]. In their analysis of (8), the preference in the interpretation in [4] is explained by zero topic assignment in [3] and preference on 'continue' transition in [4].

In summary, Walker et al.'s account of Japanese zero pronoun is based on two independent preference mechanisms: the first one is the forward center ranking, and the second is the ordered transition states. However, as Walker et al. themselves point out, each transition state between discourse segments is determined by the ordering of forward looking centers, the predictions of the theory tend to depend more largely on the forward center ranking.

3. Problems with Walker et al.
Now I would like to discuss some problems with Walker et al.'s framework. Needless to say, it has great advantages, such as relative ease of computational implementation. Moreover, I agree that accessibility of

14

discourse entities plays an important role in reference assignment, and their forward center ranking is an adequate enough approximation of accessibility of discourse entities in different grammatical categories in Japanese. However, as I mentioned before, there are cases whose interpretation process cannot possibly be explained by the accessibility factor alone. Accounts of reference assignment which are largely based on accessibility of discourse entities tend to exhibit their weaknesses when they face cases which require some pragmatic inferences, and Walker et al. is not an exception here. I will illustrate two problems they need to solve below.

3.1. Multiple topics

One of the most obvious shortcomings of Walker et al.'s approach is that it cannot handle situations where there is more than one topic in a sentence. In their framework, in order to identify zero pronouns, the backward-looking center has to be identified first. The backward-looking center, in turn, is determined by the way forward-looking centers are ranked. Therefore, the most powerful mechanism in their framework is the forward-looking center ranking shown in (5) above. However, notice that it is only useful if there is no more than one entity in each category in a sentence.

Let us concentrate on the category of topic here. In Walker et al.'s framework, the topic marker 'wa' is given a special status: the topic marker 'wa' is so powerful that the topic NP becomes the most highly ranked forward-looking center even at the onset of a discourse; in addition, once a topic NP is realised as the backward-looking center, as long as it continues to be realised as zero pronoun in subsequent discourse, it could continue to be the backward-looking center. This status of the topic NP rightly allows the possibility of multiple topics as in (9), which is very often seen in Japanese discourse:

(9)[1] *Mary to Jane wa shinyuu da.*
      and TOP best friends are

'Mary and Jane are each other's best friend.'

[2]*Senshuu no Doyoubi, Mary wa kaze*
    last week GEN Saturday   TOP cold
    o hiite nete-ita
    ACC had lying-was

'Last Saturday, Mary had a cold and was lying on the bed.'

[3]*Itsumo no youni, Jane wa*
    always GEN as   TOP
    ohiru goro denwa shita.
    noon around telephone did

'As always, Jane phoned (Mary) around noon.

[4]0 0 *eiga ni sasou tsumori datta.*
    film to invite planning to was

'(She) was planning to invite (her) to a film.'

However, cases of multiple topics cause a serious problem to Walker et al. Here, the preferred interpretation of [4] is that Jane was planning to invite Mary to a film. Walker et al. might explain this preference by saying that this is because 'Jane' is the backward-looking center in [4]. According to their framework, however, the alternative interpretation, namely, that Mary was planning to invite Jane to a film, is equally accessible, since 'Mary' could continue to be the backward-looking center in [4]. The problem is that in their current framework, Walker et al. do not provide any mechanism to choose one interpretation and discard the other.

There is another problem concerning multiplicity of entities with equal degree of accessibility. In Japanese, post-positional particles such as 'wa' and 'ga' shouldn't be used more than once in a sentence. Thus, you wouldn't come across sentences with two overtly marked topics or subjects. However, it is possible to have more than one NP with more or less equal salience in one sentence, for example, when two nouns, the first being a modifier and the second being the head noun, form a NP. A noun modifier is followed by a particle 'no', the Genitive Case particle. Some examples of NPs which contain noun modifiers are shown in (10):

(10)a. *Mary no tomodachi*
      GEN friend
      'Mary's friend'
  b. *Niwa no ki*
      garden GEN tree
      'A tree in the garden'
  c. *Tegami no henji*
      letter GEN reply
      'A reply to the letter'

Here, the first noun in each NP is the modifier. The most typical relation exhibited between two nouns combined by the particle 'no' is the 'possessive' relation, as in (10a). However, the use of 'no' is by no means restricted to that relation, as illustrated in (10b) and (10c). Now consider (11), which include a NP with this structure:

(11)(A memo written by a man, and
      addressed to his wife)
  [1]*Kooto no botan ga   toreta.*
      coat GEN button SUB came off
      'One of the buttons of (my) coat has
      come off'
  [2]0 0 *sagashi-temo,* 0 *mitsukaranakatta*
      search-although   was not found
      '(I) tried to find (it), but failed.'
  [3]*Kyou jyuu   ni* 0 0 0 *sagashite*
      today within at   find
      tsukete-hoshii.
      fix -want
      '(I) want (you) to find and fix (it)
      today.'
  [4]*Ashita* 0 0 *hitsuyou da.*
      tomorrow   need
      'Tomorrow, (I) will need (it)'

(12)[1]*Sakuban, John wa jitensha no kagi*
      last night   TOP bicycle GEN key
      o   kake-wasureta.
      ACC lock-forget
      'Last night, John forgot to use the key to
      lock his bicycle'
  [2]*Kesa,   mou* 0 *nakunatteita.*
      this morning already gone-has
      'This morning, (it) has already gone'

In (11), 'one of the buttons of my coat' is

introduced in [1] as the subject of the sentence. What is important here is that the NP introduces two conceptual entities, namely, a 'coat' and a 'button', which are equally accessible. The question is whether Walker et al.'s framework can handle cases like (11). For [2], the preferred interpretation is that the speaker wants his wife to find and fix the button today. Walker et al.'s system would successfully predict that the button is the backward-center for [2] and [3]. However, for [4], the preferred interpretation is that the speaker will need his coat tomorrow, rather than the button. I do not see how Walker et al. can explain this. Example (12) causes exactly the same problem for them. Walker et al. would predict that the preferred interpretation should be that the key has gone. This obviously is the wrong prediction. The point I would like to make here is that there are many cases in Japanese discourse where there is more than one roughly equally salient discourse entity in a sentence which subsequently become equally strong candidate referents for zero pronouns, and some mechanism of choosing the right one is needed.

3.2. Ordered preference of transition states
Now I would like to move on to a different kind of problem, which concerns their ordered preference of transition states shown in (4). Walker et al., as well as centering theorists in general, assume that when there are more than one possible overall interpretation available, the one which exhibits 'continue' transition is preferred. This is based on the assumption that maximally coherent segments are those that require less processing effort and the hearer will prefer an interpretation which requires less processing effort. However, it is not difficult to think of examples which go against their assumption. For example, look at (13) and (14):

(13)Mary [1] *Kinou* 0 *Peter to   denwa de*
      yesterday   with telephone by
      jikan o   kakunin dekita?

time ACC confirm could
'Could you confirm the time with Peter by
phone yesterday?'
    John [2] *Iya, 0 dekinakatta.*
        no   couldn't
'NO, (I) couldn't'
John [3] *0 ie ni inakatta.*
        home at present-not-was
a. (I) was not at home. [CONTINUE]
b. (He) was not at home [SHIFT]

(14)[1]*John wa joushi ni atama ga*
        TOP boss with head SUB
    agaranai.
    hold-not
'John cannot hold up his head before his boss
(i.e. John cannot help feeling ashamed of
himself in front of his boss).
    [2]*Itsumo 0 kaisha ni 0 saki ni kuru.*
        always office to earlier come
'(He) always comes to the office earlier
  (than him).'
a. (John) always comes to the office earlier
(than his boss).
b. (John's boss) always comes to the office
than (John).

Let us look at (13) first. What interests me
here is the interpretation of the zero pronoun
in [3]. Notice that both [3a] and [3b] are
equally acceptable in terms of factual
plausibility. In such case, Walker et al. would
predict that [3a], which exhibits a 'continue'
transition should be preferred. However, for
some reason, the preferred interpretation for
[3] is not [3a], but [3b]. Similarly, in (14), the
preferred interpretation for [2] is definitely
[2b]. However, Walker et al. predict that [2a]
should be preferred. Obviously, their
mechanism based on the accessibility of
transition states makes wrong predictions. In
the next section, I will consider why this is the
case.

4. A relevance-theoretic solution to problems with Walker et al.

Sperber & Wilson's relevance theory (Sperber
& Wilson 1986/95) inherits the Gricean
assumption that the hearer's goal of verbal

understanding is to find an interpretation
intended by the speaker. However, it differs
from Gricean approach in two crucial points: it
does not take the view that we have to follow
maxims, nor the view that we have to be co-
operative, to achieve successful
communication. Sperber & Wilson claim that
what makes communication achievable at all
is a fundamental mechanism built in our
cognitive system, namely, the pursuit of
relevance. This is expressed as the First, or
Cognitive, Principle of Relevance:

(15)*Cognitive Principle of Relevance*
Human Cognition tends to be geared to the
maximisation of relevance.

The notion of relevance is defined in terms
cognitive effects, i.e. some changes in the
belief system, and processing effort to obtain
such effects:

(16)*Relevance*
a. The greater the cognitive effects, the greater
the relevance;
b. The smaller the effort needed to achieve
those effects, the greater the relevance.

Cognitive effects result from the interaction of
new and old (or contextual) information in one
of the following three ways: (a) combining
with an existing assumption to yield
contextual implications; (b) strengthening an
existing assumption; (c) contradicting and
eliminating an existing assumption. Processing
effort is the mental effort needed to parse the
utterance, decide what proposition and
propositional attitude it was intended to
express, access an appropriate context, and
work out the contextual effects of the utterance
in the context. When an utterance has more
than one possible interpretation, the hearer
should look for the one which satisfies the
following conditions of optimal relevance:

(17)*Optimal relevance*
An utterance is optimally relevant to the hearer
iff:
a. it is relevant enough to be worth the hearer's

processing effort;
b. it is the most relevant one compatible with the speaker's abilities and preferences.

The Second, or Communicative Principle of Relevance, governs this search process:

(18)*Communicative Principle of Relevance*
Every utterance communicates a presumption of its own optimal relevance.

The pursuit of optimally relevant interpretation suggests a pattern of comprehension procedure the hearer should follow, which can be spelled out as (19):

(19)*Relevance-theoretic comprehension procedure*
a. consider cognitive effects in their order of accessibility (i.e. follow a path of least effort);
b. stop when the expected level of relevance is achieved.

Now let me illustrate how this comprehension procedure should work for Japanese zero pronoun resolution illustrated in (11), (12), (13) and (14). In Matsui (1995, 1998, also Wilson and Matsui 1998), I have developed the idea proposed by Wilson (1992) that in addition to the factor of accessibility of candidate referents, there is another important factor which affects the hearer's choice of referent, namely, accessibility of contextual assumptions. Accessibility of contextual assumptions becomes particularly crucial when there is more than one roughly equally accessible candidate referent, and it is the factor which is vital to solve problems with Walker et al.. In fact, the importance of contextual assumptions in reference resolution had been recognised before and various proposals were made as to how to retrieve the right context: some appeal to situationally partitioned knowledge (e.g. Sanford & Garrod 1981) and others are motivated by textual coherence (e.g. Hobbs 1979; Asher & Lascarides 1993). The account pursued here is different from any existing accounts in that it

claims that the selection of contextual assumptions is ordered in terms of *both their accessibility and likeliness to contribute towards cognitive effects of the utterance*. In other words, in relevance theory, it is assumed that these candidate referents are tested in parallel, with the one which gives quickest access to a context in which the utterance as a whole yields an acceptable overall interpretation being selected.

As a working hypothesis, let us assume that certain contextual assumptions are accessed by the hearer after the immediately preceding utterance is processed, during and after the current utterance is being processed. I have no specific claim here concerning what triggers the retrieval or the construction of certain contextual assumptions, and can go along with existing suggestions (e.g. it can be triggered by lexical information, or/and by situational knowledge). As the second working hypothesis, I would like to suggest that after having understood an utterance, the hearer tends to have, if not always, fairly accurate expectation as to what kind of cognitive effects he would like to obtain from the next utterance. Relevance theory predicts that when an utterance creates in the hearer an expectation for a specific cognitive effect to be achieved by the next utterance, other things being equal, the hearer is more likely to spend his processing effort to find an interpretation which can achieve such cognitive effect when interpreting the utterance. As a consequence, the candidate referent which is not the highest in the general accessibility ranking can become the most accessible to the hearer if the referent is expected to contribute to the interpretation he is looking for. In other words, relevance theory predicts the alteration of accessibility ranking of the candidate referents as a result of the pursuit of certain cognitive effects.

Consider examples (11) and (12) in the context discussed above. In (11), the utterance in [3] is a request to the wife to find and fix the missing button of the speaker's coat before tomorrow. Generally, if someone asks you to do something by certain time, there should be

a good reason for such a time limit, since the time limit in turn may create certain priority. Thus, it should be reasonable to assume that after hearing the utterance in [3], the question such as 'why do I have to do it today?' or 'can't it wait a little while?' occurred in the hearer's mind. If so, the utterance in [4] can readily be interpreted as the reason why he made such a request: he wants to wear the coat on the following day. The referent is assigned automatically during the process of finding expected the cognitive effects. Of course, what he needs is the coat with the button fixed, and this interpretation is only possible with the overall interpretation in which the 'coat', rather than the 'button', is the referent of the zero pronoun. The first utterance in (12), on the other hand, seems to create different kind of expectation in the hearer. After hearing that the bicycle was left unlocked, the most likely question occurs to everyone's mind is 'what happened to the bicycle?' We all know the likely consequence of the unlocked bicycle, and the hearer of (12) gets the expected contextual effects by interpreting the zero pronoun as the 'bicycle', rather than the 'key'.

In both (13) and (14), it is reasonably assumed that the hearer will have a why-question in his mind before hearing the final utterance. In the case of (13), the only possible reason why John could not confirm the time with Peter by phone was because Peter was not at home when John phoned, and the zero pronoun is resolved automatically in the process of obtaining this interpretation. I will look at (14) in more detail. The interpretation of (14) might be explained like this. After processing the first utterance, certain assumptions might become moderately accessible (but not necessarily at the conscious level) to the hearer: e.g. various assumptions about John and his boss, and more general assumptions about 'being ashamed of', e.g. that one must feel unhappy about such situation, or that one must have specific reasons for such feeling, etc. In this way, *contextual assumptions might contribute to form the hearer's anticipation about the way subsequent utterances achieve relevance.*

When the hearer interprets the second utterance in (14), further assumptions related to the event described, such as that 'workers are encouraged to come to work early', or 'bosses like their workers to arrive before them' etc., might become highly accessible. Such assumptions contribute to the hearer's search for the way the utterance could achieve cognitive effects. At the final stage of the interpretation process, the hearer finds the only way in which the second utterance in (14) might be intended to achieve relevance in a context created by the first - namely, as an explanation for why John cannot help feeling ashamed of himself in front of his boss - and the zero-pronouns are resolved automatically in the process.

Furthermore, notice that as a consequence of using the notion of accessibility of contextual assumptions, an ad hoc system such as the ordered preference in transition states in Walker et al.'s account becomes automatically unnecessary. For example, our framework could easily accommodate cases such as (13) and (14), where the referent of a zero pronoun in the current discourse is not the same as that of a zero pronoun in the preceding discourse, as well as cases where the most accessible candidate referent indeed coincides with the referent chosen on the basis of the preferred overall interpretation. Recall that the ordered transition states is created on the basis of the assumption that the hearer prefers an interpretation which exhibits 'continue' relation, because such an interpretation requires less processing effort. I suggest that this assumption is ultimately wrong for the following reason: *the preferred interpretation by the hearer is the one which provides enough cognitive effects worth his processing effort, rather than the one which merely requires less processing effort.* Walker et al. would predict that the preferred interpretation for the second sentence in (14) is that 'John always comes to his office earlier than his boss.' The question we have to ask here is: how could this interpretation possibly achieve relevance? I cannot easily see how. Relevance theory predicts that such interpretation will

never be considered when there is an alternative interpretation accessible which achieves relevance, even if the latter might require more processing effort. Notice here that the relevance theory shares the view that the intended interpretation should be the most accessible one for the hearer. However, in the framework of relevance theory, *it is accessibility of contextual assumptions, together with the accessibility of contextual referents, that determines the overall accessibility of an interpretation.* In this way, the fact that the preferred interpretation of the second sentence in (14) is more accessible overall for most of us, than the alternative interpretation, is adequately explained in this framework.

5. Summary

In this paper, I have proposed a way of complementing centering theory from a relevance-theoretic perspective. I have suggested that a model of reference assignment which appeals to the expectation based on the accessibility of candidate referents such as centering theory should ultimately accommodate some mechanism of contextual selection and pointed out that the hearer's expectation of specific contextual effects (a type of forward inferences) should be taken advantage of in creating such a mechanism. Recently, Oberlander (1998) says that 'the key lesson from the work on pronoun generation and interpretation is that we must develop a more sophisticated view of "expectation." I hope this paper makes some contribution towards that goal.

Main References

Asher, N. & Lascarides, A. 1993. Lexical disambiguation in a discourse context. *Journal of semantics* 12:69-108.

Gordon, P., Grosz, B. & Gilliom, L. 1993.

Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17:311-347.

Grosz, B., Joshi, A. & Weinstein, S. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2):203-224.

Hobbs, J. 1978. Resolving pronoun references. *Lingua* 44: 311-338.

Matsui, T. 1993. Bridging reference and the notions of topic/focus. *Lingua* 90:49-68.

Matsui, T. 1995. *Bridging and relevance.* University College London, PhD thesis.

Matsui, T. 1998. Pragmatic criteria for reference assignment: A relevance-theoretic account of the acceptability of bridging. *Pragmatics and Cognition* 6(1/2), 47-97.

Oberlander, J. 1998. Do the right thing ... but expect the unexpected. *Computational Linguistics* 24(3):501-507.

Sanford, A. & Garrod, S. 1981. *Understanding Written Language: Explorations in Comprehension Beyond Sentence.* Chichester: Wiley.

Sidner, C. 1983a. Focusing and discourse. *Discourse Processes* 6:107-130.

Sperber, D. & Wilson, D. 1986/1995. *Relevance: Communication and Cognition.* Oxford: Blackwell

Walker, M.A., Iida, M. & Cotes, S. 1994. Japanese Discourse and the Process of Centering. *Computational Linguistics* 20(2), 193-232.

Walker, M. A., Joshi, A.K. & Prince, E.F. 1998. *Centering Theory in Discourse.* Oxford: Clarendon Press.

Wilson, D. 1992. Reference and relevance. *UCL Working Papers in Linguistics* 4:167-191.

Wilson, D. & Matsui, T. 1998. Recent Approaches to bridging: Truth, coherence and relevance. *UCL Working Papers in Linguistics* 10:173-200.