Relevance Approach to Anaphora

Shin Hyun-je*
Cho Hye-sun**

1. Introduction

Anaphora\(^1\) is said to involve syntactic, semantic and pragmatic Concerning the division of labor between grammar and pragmatics regarding anaphora (Reinhart 1983; Lust 1986; Kempson 1988; Blakemore 1992), this paper argues that the contribution of pragmatics to anaphora is more crucial than has been commonly believed, and that conflicting views on the identification of anaphora in relation to the necessity of the background assumptions retrieved through inference and the status of semantic content or structural expression can be resolved within the

---

* Professor (English Department), Dongguk University.
**Instructor (English Department), Dongguk University.
1) In this paper, we adopt the term 'anaphor/anaphora' interchangeably in the general sense of a word which points back to some previous linguistic, semantic, or pragmatic object.
relevance theory introduced by Sperber & Wilson (1986).

In the following sections, we develop a pragmatic theory of anaphora within the framework of the relevance theory. We begin with an overview of the background: relevance theory and its effects, and the notion of context, which can be the whole world in relation to an utterance act, in Section 2. Then we proceed to discuss shortcomings of structural analysis of anaphoric assignment in Section 3. Finally, Section 4 presents that semantic properties of verbs, such as reciprocality or iterativeness and Caramazza's (1977) implicit causality can be regarded as an aspect of world knowledge and that its effects can be accommodated within the framework of relevance. In this theory, anaphoric assignment is determined largely by the interaction of the decoded semantic content with world knowledge assumptions about the ways in which things relate in the world when an utterance is interpreted. Utilizing world knowledge of predicates and their entailments under the relevance theory explaining and organizing their interaction, we suggest a relevance-theoretic solution to the problems that arise from the structural or semantic analysis.

2. Relevance Theory and Context

2.1. Relevance Theory

The concept of relevance which has become increasingly central in works on pragmatic inference originates from Grice's (1975) theory of conversational implicature. In his theory the maxim of relation 'be relevant' is one of the four: 'quantity,' 'quality,' 'relation,' and 'manner,' which together are taken to provide an inferential bridge between what is said and what is implicated. But recent pragmatic theories tend to view
that all maxims can be replaced by a single principle of relevance, the most significant contribution of which is made by Sperber and Wilson (1986). In contrast with Grice’s maxim of relevance which is a cooperative principle in conversation, their principle of relevance is a cost-effect principle.

Sperber and Wilson (1986) argue that this principle can account not only for semantic meaning, but also for meanings conveyed by ostension and for Gricean pragmatic implicatures (38). Their fundamental idea is that, in processing an utterance which is an overt request for attention, the hearer is looking for an optimally relevant interpretation, that is, for a correct interpretation, an utterance must (a) achieve enough effects to be worth the hearer’s attention, and (b) put the hearer to no gratuitous effort in achieving those effects (Smith and Wilson 1992: 5). Of course, it is not always the case that the interpretation the speaker intends is optimally relevant to the hearer, given inevitable mismatches of assumptions between individuals. So the criterion the hearer employs is one of ‘consistency with the presumption of optimal relevance’ a given interpretation of an utterance meets this criterion if and only if the speaker of that utterance could rationally have expected that interpretation of it to be optimally relevant to the hearer (Sperber and Wilson 1986). According to Kempson (1988), anaphoric interpretation process well displays the effects of the principle of relevance.

2.2. Context

In a broad sense, 'context' is one of those linguistic terms which refers to the relevant elements to the surrounding linguistic or nonlinguistic structures in relation to an utterance under consideration. In other words context is the ensemble of locally and globally operating conditional
relevances in their forward and backward directions. The context is, therefore, not uniquely determined, but chosen from within a range of available contexts. Each individual has at his or her disposal a set of assumptions, including new assumptions and those already stored in memory, the latter acting as a context to the former. In order to identify the context, each individual needs to (a) access previous assumptions recently used and (b) retrieve his or her encyclopedic entries from the stored memory.

Although these two procedures may look different and can be analytically separated, they interact with each other in that encyclopedic entries can be partly inferred from the previous assumptions. A gap between (a) and (b) is filled not by more coding, but by inferencing. The relevance-theoretic account places heavy emphasis upon the hearer's encyclopedic information. As Blakemore has indicated, it is these background assumptions which form the context of the utterance. The following example illustrates this point:

(1) I walked into John's kitchen. The windows were spotless.

In (1), the correct interpretation is not recovered on the basis of linguistic rule, but on the basis of contextual, that is, extra-linguistic information made available by the interpretation of the first segment in the sequence in accordance with the principle of relevance. A hearer has in general access to an enormous amount of background information, and any of this could be used in the interpretation of an utterance.

According to Sperber & Wilson's claim that utterances automatically create expectations of relevance, since the content of all anaphoric expressions provides us with a guarantee that an antecedent is immediately available, we predict that where no antecedent is provided by the explicit content of the discourse, it must be provided by additional
premises as part of the context. This is the situation in the bridging cross-reference cases (Kempson 1988). The hearer deduces that the speaker must have been assuming the accessibility of the additional premise, 'John’s kitchen has windows.' So the windows in the second utterance may be retrieved from the encyclopedic entry of the concept of kitchen. That is, the interpretation of (1) depends on the hearer’s encyclopedic information that a kitchen has windows, under appropriate conditions consistent with the principle of relevance. Thus, the encyclopedic information could be part of the required contextual information. We assume here that although extra processing effort has been required of the windows, this effort is not so excessive because of the easy accessibility of the hearer’s encyclopedic knowledge. A kitchen has windows.

Let us consider the second one:

(2) John took the train to Pusan yesterday.
   The trip took 3 hours.

The most accessible context for the identification of definite NP anaphor is derived from the immediately preceding sentence. However, in (2) since there is no explicit mention of trip in the preceding sentence, the hearer has to resort to something more than the representation assigned to the preceding linguistic expression. Its denotation is inferred from the whole proposition or event. Thus, for example, the antecedent for the trip is the trip associated by the preceding proposition John took the train to Pusan. It seems to be encyclopedic knowledge about going somewhere by train and traveling. This knowledge is indispensable for identification of the intended reference. It must be established by the hearer in his mental representation using clues given the explicit preceding discourse. In other words, definite NP anaphor the trip is identified by the
interaction between the immediate preceding context and encyclopedic information about *the trip*.

3. The Problems of Structural Analysis

Here we would like to show that the analysis along the line, mentioned above fails to resolve shortcomings of structural analysis. Consider the following example taken from Hirst (1981).

(3) John drank the wine that was on the table.
   a. It was brown and round.
   b. It was delicious.
   c. It had been there since the morning.

The unacceptability of (3a) cannot be attributed to the fact that the referent of the pronoun *it* of (3a) is ambiguous between *the wine* and *the table*, because (3b), in which the antecedent of *it* would be equally ambiguous, is acceptable, while (3a) is unacceptable despite the fact that the hearer can access the hypothesis from encyclopedic knowledge—that is, a table can be brown and round.

Why is the correct referral the one that links *it* to the wine that John drank and not that links *it* to the table? Hirst suggests that a purely syntactic explanation based on depth of embedding would be simpler. Following Hirst, the infelicity of the coreferential relation between the antecedent and the pronoun in (3a) lies in the embedded syntactic position of the antecedent. In (3b), on the other hand, a less embedded NP *the wine* is easy to interpret as an antecedent of the pronoun. Erku & Gundel (1987) claim that an anaphoric expression will refer to the
item that is currently in topic, which is most likely to be introduced in
the verb phrase of a main clause (in direct object position, if there is
one). At first sight, it might seem that the syntactic position of an NP
can tell the hearer what is likely to become important for the rest of the
discourse in order to avoid infelicitous references like this. However,
there are additional data with which these purely syntactic explanations
are inconsistent.

For example, (4) is syntactically similar to (3). However, in (4) the
object of a preposition Mary inside a relative clause, can serve as the
antecedent of she.

(4) John drank the wine that he was going to give to Mary.

She was disappointed.

Although (3) and (4) have the same syntactic structure, Mary in (4) is
more deeply embedded than the table in (3). Thus, if the notion of topic
is based entirely on syntactic embedding, it would be more difficult for
Mary to be the antecedent of the pronoun she than for table to be the
antecedent of the pronoun it in (3). If, as Erku and Gundel claim, there
is only one local topic information in each sentence, which should be the
direct object in (4), this example should be judged odd, because the
pronoun she in the second sentence does not refer back to the predicted
topic the wine. Hence, this example demonstrates that Hirst’s and Erku
and Gundel’s accounts fall short. Those examples (3a) and (4) also
suggest that the difference in acceptability between the two cannot be
explained by purely structural analysis.

Now we would like to suggest a relevance-theoretic solution to the
problems that arise from those accounts. By the theory of relevance, the
hearer can assume that the speaker would not have drawn the hearer’s
attention to a constituent, unless the speaker wanted to talk about it.
and unless it became the most natural topic for the hearer's next utterance. In other words, the next utterance, if it is to be cooperative, should be about some relevant aspect of the constituent.

Consider the example (3) again. The hearer can assume that the speaker intended to direct his attention to some aspect of a constituent, namely, *the wine*, in which case (3b) is acceptable, while (3a) is odd. Most hearers would interpret *it* in (3a) as *wine* rather than as *table*. Recall that the hearer's aim in the reference assignment is to construct a proposition which is optimally relevant in a way that the speaker could have manifestly foreseen. There might well be a context in which it would be relevant that *table* is referred. However, given the assumptions about what people expect when they drink something, it is easier, on the basis of what has gone before, to construct a context in which it would be relevant to expect to hear something about drinking. Hence, (3a) is odd, because after hearing the first sentence, the hearer is expecting to hear something about drinking wine, not about table. More strictly speaking, the reason that (3a) is odd seems to be that in the context up to the expression, *the wine* introduced by (3), it implies not a bottle with wine, but drinkable liquid and so the expression *round* is not appropriate to an appearance for liquid, which is established as given by encyclopedic knowledge. It can therefore be said that *it* in (3b) refers to the referent *the wine tasted by John* rather than only to a single NP *the wine*. In contrast to (3b), in (3c) *it* refers only to the existence of *the wine*.

Now consider the example (4) again. Drinking the wine that John was going to give Mary is a sufficient cause for her disappointment. We know this from the fact that if the second utterance in (4) is prefaced by *so*, it will be immediately clear whether the proposition expressed is intended to be conclusion, and the hearer will be encouraged to access appropriate contexts to make the arguments go through. As Blakemore (1987) points
out, such contexts contribute to relevance by guiding the hearer towards the intended contextual effects, hence reducing the overall effort required. The speaker who connects the two utterances in (4) must intend an interpretation on which the effort of processing them jointly is justified by extra effects. Such an interpretation would be achieved if it was relevant to know why Mary was disappointed. The main function of contextual assumptions in utterance interpretation is to enable the hearer to establish coherence relations and optimally relevant interpretation.

Here we will show how restrictions on anaphora in coordinate structures can be related to the principle of relevance. To account for this, we propose that coordinate NPs are treated as pragmatic units in that both conjuncts have to agree in givenness or topic-hood. This is consistent with the co-ordination principle (Lyons 1977:405). Erteschik-Shir and Lappin (1979) state that one conjunct cannot be interpreted as dominant (focus) without the other. If the same is true for givenness, this would mean that both NPs must be given or new. The use of an anaphoric reference means that the referent or the set to which the referent belongs is given. Thus, if a pronoun is coordinated with a non-pronoun, the resulting coordinate NP should be unacceptable unless the non-pronoun is also given by the context determined by the principle of relevance.

This proposal explains why example (5) seems better than the other example of a pronoun with an indefinite NP as illustrated in (6).

(5) John went to the hospital. ??He and a doctor talked there.

Although (5) is not fully acceptable, it is better than a sentence with an unrelated full NP like in (6).

(6) John went to the hospital. ??He and a judge talked there.
Contextual information is stored under a concept (e.g., hospital), and the hearer is given access to the information stored under it when the concept is presented to him for processing. That is, if the concept hospital is presented, the hearer is given access to further information that there are doctors there, not judges. Thus, it may be the case that doctors are somewhat accessed from hospital in (5). This seems to be an example of Blakemore's (1992) 'stereotypical knowledge.' Bosch (1985) considers this as the control of attention by expectations that are arisen by what has been perceived already. Fundamental notions about this kind of control are sketched in papers on 'frame' (Hawkins 1978).

Thus restrictions on coordinate anaphora are due to a pragmatic requirement that elements of a coordinate structure must agree in givenness. Specifically, one element cannot be given when the other is not. And this may be a consequence of a general fact that coordinated items must be treated as a pragmatic unit. Accordingly, because the acceptability of the coordinate anaphora examples varies with the discourse context, as shown in (5) and (6), coordinate anaphora can only be analyzed as cognitive pragmatic phenomenon consistent with the principle of relevance. The hearer interprets utterances by computing its cognitive effects in a context based on the preceding utterance and his encyclopedic knowledge of what it is to be a hospital and a doctor, in order to establish adequate contextual effects, which can minimize the processing effort.

We have shown that the givenness helps the hearer determine the antecedent of anaphora, when a concept becomes given by the process of the context constructions controlled by the principle of relevance. From the point of view of communication, it might be said that items become given when they may be important by the hearer in the rest of the discourse. This potential importance can hardly be signalled by syntactic factors, but rather the hearer can infer it on the basis of relevance-
theoretic approach plus a given concept that is assumed by the speaker to be in the hearer’s consciousness.

4. World Knowledge of Predicates

4.1. Split Antecedents

This section is concerned with split antecedent phenomena. If more than one nominal serve as antecedents of an anaphoric expression, such antecedents are called ‘split antecedents.’ Let us consider the following example:

(7) John met Bill at a coffee shop. They had a chat.
(8) John found a good piano teacher.

They are very fond of Mozart.

In (7), the split NPs\(^2\) John and Bill can together serve as antecedents to the plural anaphora they. In (8), on the other hand, regardless of the same structural condition as in (7), the split NPs John and a good piano teacher are not likely to serve as antecedents of they in (8). But if the expression both is added to (8), then the sentence becomes acceptable, as shown below:

(9) John found a good piano teacher.

They are both very fond of Mozart

---

2) There is also a case in which split antecedents and their anaphors occur together within a single sentence as in (i):
(i) John\(_i\) talked to Bill\(_j\) about their\(_i, j\) problem.

The subscripts \(i\) and \(j\) are to be interpreted as variables, under the tacit condition that \(i \neq j\).
They in (9) seems to be licensed in part as an effect of both. It does not affect the truth-conditional content of utterances in which it occurs. The occurrence of both with a plural pronoun they might serve to reinforce the plural grouping, which again enhances the activation of contextually present information about the antecedents John and a good piano teacher. In other words, the use of both instructs the hearer to establish an inferential connection between the preceding two nouns and they. Both makes an essential contribution to the comprehension process in the similar way as connectives do according to Blakemore (1987, 1992). This contribution helps to minimize the processing effort. From the examples seen here, it seems that a simple structural or semantic characterization would be quite difficult. We need a general way to characterize all of the possible antecedents. For this purpose we shall take both semantic and pragmatic aspects into consideration to account for split antecedent phenomena.

According to Sanford and Lockhart’s (1990:366), what licenses plural reference is the possibility of mapping atomic individuals into a common role-slot, namely, two split antecedents’ playing the same role in an acting scene, in some interpretive structure. In contrast to Sanford and Lockhart, we propose that plural pronoun is interpreted by grouping individual elements introduced separately into encyclopedic knowledge and then by forming some sort of conceptual link between the individual

3) Sanford and Lockhart (1990:367) explain the mapping process within a main verb-schema as in the following example (i):

(i) John asked Mary out.

The sentence (i) might be represented by mapping in which John is an agent and Mary is a patient in the verb schema. Such mappings are assumed to be made at a number of levels. For example, if someone knew that John and Mary always talk a lot when they go out, then that inference might be made in example (i), leading to a further set of mappings in which both John and Mary map into the same role-slot (both are agents in a talking schema).
elements. Since this process may be derived from the hearer existing representation of the world, an inference system may play a role in the conceptualization.

In (7), for example, we know that John and Bill met in order to carry out joint activities, that makes us have the background knowledge that John and Bill met to do something together in (7). Kuno (1987:242) also notices that the verb meet implies the marked nature of the reciprocal verb. That is, the two characters John and Bill play the same role in a meeting scene. So they are considered to act together in this scene. A plural pronoun is therefore possible since the two individual elements are expected to do the same action, and then, a conceptual link can be formed between them.

In contrast, in (8), only John plays a dominant role as an agent in the action of finding, while a good piano teacher does not. It is John, not a piano teacher that is crucial to the action of finding. The two persons John and a good piano teacher therefore play different roles in the conception suggested by the sentence (8). Hence it requires more effort to process (8). This interpretation would not be satisfactory on the effort side, and so inconsistent with the principle of relevance. It seems that they are inferred to continually play different roles in the next utterance. Our inference of this based on the semantic aspect of the verb find hinders us from putting the two characters in the cooperation.

In short, if split antecedents take the same action in an acting scene, then each individual referred to by split antecedent is inferred to play an equal role and so plural anaphora is possible. On the other hand, if they do not, then plural anaphora is not easily available. The hearer interprets the speaker’s utterance by computing its cognitive effects in a context based on the preceding discourse, and his knowledge of the meaning of verbs and the split antecedents.
4.2. Semantic Bias of Verbs

A similar claim regarding the effect of the semantic properties of verbs on anaphora in clause boundaries is suggested by Caramazza, Grober, Gravey and Yates (1977). They argue that implicit causality as a property of verbs is responsible for biasing antecedent assignment from subordinate clauses to main clauses which denote the action or state. They present sentences containing verbs of type 1: such as win, telephone, sell, etc., and type 2: such as envy, fear, etc.. The sentences are built so that the subordinate clause could or could not be compatible with the semantic bias produced by the verb of the main clause. A pronoun in a subordinate clause, for example, is coreferential with the subject of the main clause containing type 1 verbs like win, sell, telephone, etc.

Consider sentences containing the verb telephone, taken from Caramazza et al. (1977).

(10) a. John\textsubscript{i} telephoned Bill because he\textsubscript{i} wanted some information.
    b. John telephoned Bill\textsubscript{i} because he\textsubscript{i} withheld some information.

The sentence containing the verb telephone as in (10a) typically makes the anaphora he refer to the main subject. Since the subordinate clause of a sentence like (10a) is consistent with the bias introduced by the verb telephone, the appropriate antecedent John is easily and naturally selected for anaphora he. This cannot be the result of syntactic factors because in a syntactically similar sentence such as (10b), he would probably be understood as referring to Bill. One way of accounting for

4) In a semantic-analytic perspective, the cause of the action of type 1 verbs is attributed to the agent of the verb and the cause of the action of type 2 verbs is attributed to the recipient of the verbs as the following example shows. Example: I often envy the writer, because he works in a university.
this predisposition is to attribute it to a factor called implicit causality, in which anaphora is resolved so as to keep role assignments (subject, object, etc.) fixed. Since in the sentence (10b), the interpretation of the subordinate clause opposes the semantic bias produced by the verb of its main clause, not the subject but the object Bill in the main clause is selected as the antecedent to anaphora he. As we have seen, thus, implicit causality as the semantic property of verbs is an important determinant of pronoun assignment in the subordinate clause.

A semantic property of verbs by Caramazza et al., may be regarded as an aspect of world knowledge and its effects can readily be accommodated within the framework of relevance. World knowledge includes information about the implicit causality of the matrix verb. Use of the verb telephone, for example, typically implies that the cause of the action (of telephoning) is to be located in the telephoner. However, when it is not compatible with the bias of the verb as in (10b), world-knowledge-based inferences are needed to establish the referent of a pronoun. So proper anaphoric assignment requires full interpretation of the context of the subordinate clause. It denotes a kind of event that our world knowledge tells us is prompted by something about the telephoner or telephonee. Therefore, reference assignments depend on the person who gives a reason for such an action to the event.

World knowledge may become part of the context for the interpretation of utterances. It can also give the hearer immediate access to contextual assumptions. Using them, the hearer can derive contextual effects, which are said to be controlled by the principle of relevance.

4.3. Iterative Aspect

If we take semantic properties of expressions as the basis of linguistic
content, we are forced to give up providing a unitary account of anaphora.

Let us consider the following examples, taken from Stirling (1985: 74).

(11) a. Each day, Mary wrote a letter to her sister.

    They were chatty letters.

b. Everytime John went abroad he bought a souvenir.

    They just lie around the house.

In (11), from the fact that plural anaphors are related to singular antecedents, it may be predicted that the plural anaphora they inherently denotes a set of accumulated individuals. On the basis of the truth conditions of the utterance and reference assignment, the hearer of (11) would be able to recover only a trivially true proposition.

On the assumption that the speaker is presenting information which yields adequate contextual effects, the hearer will recover a much more completely specified proposition. That is, in (11a), the assumption of the first utterance, via the interpretation of each day and writing a letter, invokes the representation of a plural NP letters to serve as an antecedent to the anaphora they. In other words, the relevant assumption of writing a letter each day would be the basis for the representation of letters. Similarly, in (11b) the interpretation of the first utterance, via the compositional interaction of the distributer every and the meaning that a singular NP a souvenir has in isolation, invokes the representation of plural NP souvenirs to serve as the antecedent to the anaphora they. That is, the assumptions of everytime going abroad and buying a souvenir the hearer might deduce would be the basis for the representation of souvenirs.

A hearer interprets the pronoun in the expectation that it will interact with the preceding assumptions to yield what Sperber & Wilson (1986)
call a contextual effect. The hearer will assume that the speaker has produced an utterance that is worthy of his attention. That is, he will assume that the speaker has been relevant, and that utterances automatically create expectations of relevance. A gap between the semantic representations of sentences and the thoughts communicated by utterances, Sperber & Wilson (1986: 9) claim, is filled not by more coding, but by inference. In interpreting the utterances the hearer infers the plural anaphors which are not linguistically encoded in the first utterances. It can be said that the drawing of inference, in which contextual assumptions maximize their relevance, should be added to the semantic representation. So the example (11) suggests the contextual implication is derivable only when the discourse context and the truth-conditional content are considered in combination in the relevance theory. In other words, the example (11) is the result of semantic representation in accordance with the hearer's aim of recovering an interpretation consistent with the principle of relevance.

5. Conclusion

We have tried to show in this paper that on some earlier approaches using semantic or structural considerations, and givenness, we cannot see adequate accounts of how anaphoric expressions are recognized and interpreted, if these factors play crucial roles in the interpretation of anaphoric reference. It is clearly wrong to pay attention only to these factors for anaphoric interpretation. Rather it is necessary to pay serious attention to contextual factors and inferential processes. A purely syntactic explanation of the use of anaphora based on depth embedding in a subordinate structure should be replaced by pragmatically based
approach consistent with the principle of relevance. Thus we considered that entailments or implicit causality as a semantic property of verbs may interact with world knowledge assumptions about the way in which things relate in the world for the interpretation of an utterance. World knowledge can give the hearer immediate access to contextual assumptions. The contextual effects can readily be accommodated within the framework of relevance.

We have also noted the semantic/pragmatic interaction and its possible relation to a theory of how anaphora is interpreted. The choice of the antecedent of anaphora seems to be depend on pragmatic factors, while keeping semantic considerations to a bare minimum. The gap between the pragmatic factors and semantic considerations can be bridged by the principle of relevance. This involves an interaction of the decoded semantic content with general knowledge assumptions about the world. From this, we have shown how, given a fairly simple characterization of the meaning of predicates, relevance theory provides a natural explanation for some aspects of meaning which arise in use. The relevance-theoretic pragmatic approach advocated here does embrace a much wider range of examples than has been commonly believed.
References


Kuno, S. 1987 *Functional Syntax*. Chicago: Chicago University Press,
Lust, B. 1986 *Studies in the Acquisition of Anaphora I: Defining the Constraints.* D. Reidel Publishing Co


Stirling, L. 1985 Distributives, Quantifiers and a Multiplicity of Events. *European Chapter of the Association for Computational Linguistics.* 16-24