

Toward a Functional Theory of Anaphora*

Dong-Whee Yang

1. Introduction

One of the most recalcitrant areas in generative-transformational grammar has been anaphora. It is mainly due to the fact that the theory of generative-transformational grammar has been strictly sentence-bound and constituent-structure-dependent whereas anaphora phenomena are typically discourse-oriented and characteristically dependent on communicative functions rather than constituent structures. Recently linguists like Kuno (1975) and Bolinger (1977) have proposed a discourse-oriented functional approach to anaphora. This paper is to contribute to their attempts by proposing six additional hypotheses for the functional approach to anaphora.

2. Sentential vs. Discourse/Contextual Anaphora

The first hypothesis we propose is (1).

- (1) *Hypothesis 1*: Sentential anaphora and discourse/contextual anaphora, or linguistic anaphora and nonlinguistic anaphora, are one and the same phenomenon.

Hypothesis 1 implies that sentential anaphora and discourse/contextual anaphora have to be accounted for by the same rules or theory of anaphora since they are the same phenomena. According to this hypothesis, all the earlier studies on anaphora that treat sentential anaphora as somehow distinct from discourse/contextual anaphora and try to account for the former as separate from the latter are misguided.

Motivations for Hypothesis 1 are as follows. Consider (2).

- (2) a. *John* wore *hĭs* coat.¹
b. John wore *his* coat.

The only difference between (2a) and (2b) is that *his* in (2a) is unstressed whereas *his* in (2b) is stressed, as indicated by the stress marks. In (2a) *John* and *his* are understood as coreferential, but in (2b) *John* and *his* are never understood as coreferential, *his* being coreferential with someone in the discourse or nonlinguistic context.

* This paper was presented at the Spring Meeting of the English Literary Society of Korea (in Taegu), June 3, 1978. I am grateful to the audience for their comments on this paper. The term 'functional' in this paper covers anything 'nonsyntactic' or 'non-syntax-centered', the term 'functional theory' being opposed to 'traditional sentence-bound syntactic theory'.

¹ Italicized words or phrases are in a coreferential or anaphoric relation. "˘" indicates a stress whereas "˘" does no stress.

Now compare (2) with (3).

- (3) a. John hit *Bill* and then Mary hit *him*.
 b. *John* hit Bill and then Mary hit *him*.

Again, in (3a) *him* is unstressed whereas in (3b) *him* is stressed. Note that in (3a) *him* is coreferential with *Bill* whereas in (3b) *him* is coreferential with *John*.

(2) and (3) show that when pronouns are stressed they do not maintain the normal coreference relations, i.e. they refer to some other NP than the one they would normally refer to if they are unstressed. Thus, in (3b) the stressed *him* refers to *John* rather than to *Bill*, and in (2b) the stressed *his* refers to someone that is not mentioned in (2b) since there is no NP other than *John* within the sentence. Since it is most natural to assume that the coreference relations of the stressed *his* in (2b) and the stressed *him* in (3b) be accounted for by the same type of rules or theory, sentential anaphora and discourse/contextual anaphora have to be treated as the same phenomena in our grammar.

For a more obvious piece of motivation for Hypothesis 1, compare (4a) and (4b).

- (4) a. *My brother* is a doctor, and *he* says that your hair will fall out if you eat that.
 b. *My brother* is a doctor. *He* says that your hair will fall out if you eat that.

In both (4a) and (4b) *he* is coreferential with *my brother*²; but the two NP's are within a single sentence in (4a) whereas they are in two separate sentences in (4b). The coreference relation in (4a) may be stated or captured within the framework of the current theory of generative grammar, but that in (4b) may not. However, it would be most unnatural that the two coreference relations in (4a, b) should be characterized in two different ways or theories. Obviously any linguistic analysis or explanation for the anaphoric relations in (4a, b) has to be uniform.

Furthermore, consider (5), a piece of dialogue between A and B.

- (5) A: *My brother* is a doctor.
 B: *He* says that my hair will fall out if I eat that.

Note that the nature of the anaphoric relation between *my brother* and *he* in (5) is essentially the same as that in (4a) and (4b). Note also that the anaphoric relation between *my brother* and *he* in (5) is across a speaker-boundary while that in (4b) is across a sentence-boundary.

Anaphoric relations like that between *my brother* and *he* in (4) and (5) may be across the linguistic vs. nonlinguistic boundary, as we see in (6).

² In (4a, b) *he* may refer to some other man mentioned or suggested in the discourse or context involved.

(6) A: (A points to *his brother*.)³

B: *He* says that my hair will fall out if I eat that.

In (6) the antecedent of *he* can be assumed to be suggested by A's gesture or pointing to his brother. Here we see that anaphoric relations may involve nonlinguistic contexts. If we assume that it is most desirable to characterize the anaphoric relations between *he* and *my brother* or A's brother in (4a, b), (5) and (6) in a uniform way in our grammar, then sentential anaphora as in (4a), discourse anaphora as in (5) and contextual anaphora as in (6) have to be treated as essentially the same phenomena.

Hankamar and Sag (1976) point out a difference between discourse anaphora and contextual anaphora, as illustrated in (7) and (8).

(7) A: I am going to *pass this ball through this hoop*.

B: I don't think you can ϕ .

(8) A: (A is attempting to *pass a 9-inch ball through a 6-inch hoop*.)

B: #I don't think you can ϕ .⁴

I don't think you can *do it*.

The null VP anaphora, they argue, is possible with a linguistic antecedent as in (7), but it is not possible with a nonlinguistic or contextual antecedent as in (8), i.e., only the *do it* VP anaphora is possible in the latter.

Their claim, however, is not entirely correct. As Schachter (1977) argues, the null VP anaphora is possible with a nonlinguistic or contextual antecedent in cases like (9) and (10). The antecedents of all the three null VP anaphors in (9) and (10) are nonlinguistic or contextual.

(9) John: (John tries to *kiss Mary*.)

Mary: John, you mustn't ϕ .

(10) John: (John comes to the table where Mary is sitting, makes as if to *take one of the spare chairs there*.) May I ϕ ?

Mary: Please do ϕ .

Furthermore, Schachter (1977) offers an explanation for why the null VP anaphora is impossible in cases like (8) whereas it is possible in cases like (7), (9) and (10). Null VP anaphors, he argues, require more specific information for determination of their antecedents than *do it* VP anaphors do because the referential range of the former is wider than that of the latter, and in the cases of (9) and (10) the null VP anaphors are possible since their antecedents are clear from the contexts. According to Schachter (1977), then, null VP anaphora is impossible in (8) not because the antecedent is nonlinguistic or contextual but because the information available for determination of the

³ The statement in the parenthesis indicates actions or gestures performed by the interlocutor.

⁴ # indicates that the utterance or sentence so marked is incompatible with the indicated context, e.g., A's attempt to pass the ball through the hoop in (8). ϕ indicates a null anaphor.

antecedent is not enough for a nonlinguistic anaphora. Therefore, null VP anaphora does not warrant a critical distinction between discourse anaphora and contextual anaphora.

Williams (1977) suggests another difference between discourse anaphora and contextual anaphora, as illustrated in (11) and (12).

- (11) A: I'm going to *pass this ball through this hoop*.
 B: I don't think you can *do so*.
 (12) A: (A is attempting to *pass a 9-inch ball through a 6-inch hoop*.)
 B: #I don't think you can *do so*.
 I don't think you can *do it*.

The *do so* VP anaphora is possible with a linguistic antecedent as in (11), but it is impossible with a nonlinguistic antecedent as in (12). Williams (1977) claims that pronouns dominated by NP, such as *it*, *he*, etc., may refer to a contextual antecedent whereas those not dominated by NP, such as *so*, *such*, etc., may not, as evidenced in (12).

His claim, however, does not necessarily lead to recognition of a categorical distinction between discourse anaphora and contextual anaphora. The phenomena of (12) may as well be accounted for by the difference in degree of anaphoricity between the NP-dominated anaphors like *it*, *he*, etc., and the non-NP-dominated anaphors like *so*, *such*, etc., irrespective of the distinction between discourse anaphora and contextual anaphora.⁵ In fact, his claim does not apply to Korean cases. In Korean we can say (13) to someone doing what we think he should not do.

- (13) *kilə-ci ma!*
 do so don't
 'Don't do so!'

kilə 'so' is definitely not dominated by NP and yet it may refer to a contextual antecedent as in the above case.

Our second hypothesis is (14).

- (14) *Hypothesis 2*: Sentential anaphora is a special case of discourse/contextual anaphora.

Hypothesis 2 implies that anaphora is essentially discourse/contextual phenomena and that sentential anaphora should be viewed in that perspective.

For a motivation of Hypothesis 2, consider the following piece of dialogue (15).

- (15) A: *John* wants to see you.
 B: *Bill* wants to see you, too.
 C: If *Peter* shows up, please let *him* see me.

⁵ Refer to Lakoff (1968) and Cole (1974) for different degrees of anaphoricity of various anaphors.

Suppose that A and B are secretaries of the employer C, and that the piece of dialogue (15) occurred when C entered his office. Depending on the context of the dialogue and C's gesture in uttering the sentence, the antecedent of the pronoun *him* in C's utterance could be either *John* or *Bill* or *Peter*. The fact that any preceding NP representing a human male can be the antecedent of the pronoun *him* in (15) would be best characterized as a discourse/contextual phenomenon rather than a sentential phenomenon.

Furthermore some seemingly sentential anaphora phenomena cannot be accounted for unless considered in discourse/contextual perspective, as we see in (16) and (17).

(16) A: Tell me about *John*.

B: Although she dislikes *him*, she is still seeing *John*.

(17) A: Tell me about *Mary*.

B: # Although she dislikes *him*, she is still seeing *John*.

John and *him* in B's utterance are coreferential in the context of (16), but not in that of (17). Furthermore, *John* and *him* in B's utterance become coreferential even in the context of (17), if the latter is embedded in a broader proper discourse or context, as we see in (18).

(18) A: How is *John* doing these days? Is *he* still going out with many girl friends?

B: Yes, *he* hasn't changed a bit.

A: How about *Jane*? Is she still seeing *John*?

B: Yes, she is still devoted to *him*.

A: Tell me about *Mary*.

B: Although she dislikes *him*, she is still seeing *John*.

As Kuno (1975) points out, *John* and *him* are coreferential in B's last utterance in (18) despite the preceding A's utterance, because *John* is the topic of the whole discourse (18). That is, in (17) the topic is *Mary*; therefore, *him* cannot be coreferential with *John* in B's utterance. But in (18) the topic is *John*, so *him* can be coreferential with *John* in B's last utterance. Apparently the topic of the whole discourse (18) overrides that of a sub-discourse, i.e., the last A and B's utterances in (18), in determining coreferentiality of *John* and *him* in B's last utterance. Thus, for cases like this, sentential anaphora has to be viewed in discourse/contextual perspective. In other words, the consideration of discourse/contextual anaphora should be prior to that of sentential anaphora, or sentential anaphora should rather be considered as a special case of discourse/contextual anaphora.

3. Marked vs. Unmarked Anaphora

Our third hypothesis is (19).

(19) *Hypothesis 3*: The unmarked anaphora is always forward.

This hypothesis is intuitively highly plausible. In fact, in all the examples we have discussed so far in this paper the anaphoric relation is forward, i.e., the antecedent precedes the anaphor, irrespective of whether they are sentential, discourse, or contextual anaphora. For example, even the seemingly backward sentential anaphora of (16) has been shown above to be a case of forward discourse anaphora,—that is, the antecedent of *him* in (16) is not the following *John* of (16B) but the preceding *John* of (16A).⁶ Indeed if the preceding NP is not *John* in such a case then the *him* and the following *John* cannot be coreferential as we see in (17). The antecedent of *him* in B's last utterance in (18) has also been shown to be the topic of the whole discourse, *John*, rather than the following *John* in the B's last utterance, hence again a case of forward discourse anaphora. In fact, Delisle (1974) argues that most of the examples of backward pronominalization in such works as Langacker (1969) and Ross (1969) are extremely awkward or ungrammatical to him in the intended sense unless preceded by a sentence that contains the possible discourse antecedent of the pronoun. That is, he claims, most cases of backward sentential pronominalization are in fact those of forward discourse pronominalization.

Kuno (1975) points out that (20), in which *Jane* and *her* are supposed to be coreferential, is acceptable if *Jane* is sufficiently 'up' in the minds of both the speaker and the hearer.

(20) Hey, Gus—would you marry *her* if *Jane* asked you to?

The fact that *Jane* has to be topmost in the minds of the speaker and hearer when the sentence is uttered in order for it to be acceptable implies that the antecedent of *her* cannot be the following NP *Jane*. I claim that the antecedent of *her* in (20) is the *Jane* in the minds of the speaker and hearer and that it is a case of *forward* contextual anaphora since the *Jane* is topmost in the minds of the speaker and hearer right before the speaker utters (20).

Now the question is: how should we account for the anaphoric relation between *him* and *John* within (16B) or that between *her* and *Jane* within (20)? In other words, why do we return to a full NP once we have already pronominalized the preceding coreferential NP by virtue of its discourse or contextual antecedent? In fact we do not have to return to a full coreferential NP in cases like (16B) and (20). Indeed (21) can be substituted for (16B) in (16), and (22) for (20) in the very same context.

(21) Although she dislikes *him*, she is still seeing *him*.

(22) Hey, Gus—would you marry *her* if *she* asked you to?

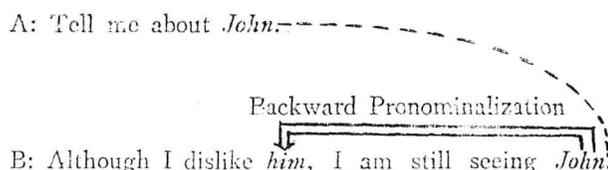
I claim that the full coreferential NP's in cases like (16B) and (20) are results of

⁶ When the pronominalizing NP follows the pronoun, the former should be named as 'postcedent' rather than 'antecedent'. But in this paper I will use the term 'antecedent' for such backward pronominalizing NP's too.

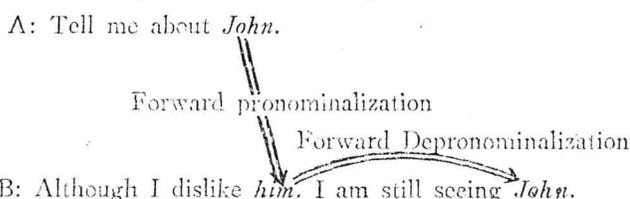
a process or rule of what I call NP Reidentification, which reidentifies the once-pronominalized-NP with a full NP later under certain conditions. I further claim that the NP Reidentification process does not relate the preceding pronoun and the following full NP in any direct way without any reference to the preceding discourse or contextual antecedent of the pronoun, in cases like (16) and (20), and that the following full NP does not pronominalize the preceding pronoun backwards, nor the preceding pronoun depronominizes the following full NP forwards, but the discourse or contextual antecedent is simply reidentified as a full NP after having once been pronominalized.

We can schematically represent the NP Reidentification approach in comparison with the other two possible approaches, as we see in the following.

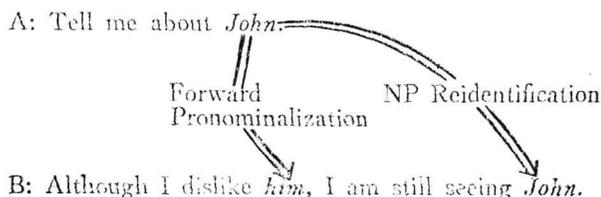
(23) a. Backward Pronominalization Approach



b. Forward Depronominalization Approach



c. NP Reidentification Approach



(23a) is the approach proposed by Kuno (1972:302), who explicitly states:

Backward pronominalization is possible in English only when the rightmost of the two coreferential noun phrases represents old predictable information.

The dotted line in (23a) indicates that *John* in B's utterance is old predictable information due to the presence of *John* in A's utterance. The problem with this approach is that it is impossible to account for the anaphoric relation in (23a) and that in (24) in a uniform way.

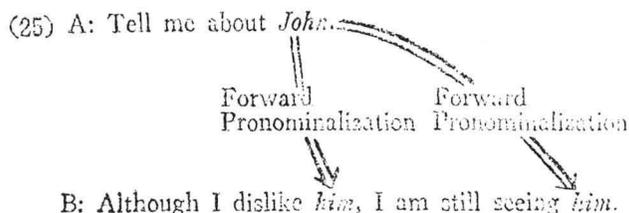
(24) A: Tell me about *John*.

B: Although I dislike *him*, I am still seeing *him*.

Note that backward pronominalization can never be assumed to be involved in (24B), whereas it is assumed to be so in (23aB) by the Backward Pronominalization Approach.

The first half of both Forward Depronominalization Approach and NP Reidentification Approach is the same Forward Pronominalization process relating *him* in B's utterance to *John* in the preceding A's utterance. This forward discourse pronominalization process is suggested by Delisle (1973) for cases like (16) and (23). Delisle (1973), however, is not explicit about the latter half of the both approaches (23b, c).

We prefer NP Reidentification Approach to Forward Depronominalization Approach for the following reasons. First, the process of Depronominalization, or blocking further pronominalization of an once-pronominalized NP, is a rather unnatural or unlikely process whereas that of NP Reidentification, which repeats, rather than pronominalizes, a preceding NP under certain conditions is a rather likely process. Second, NP Reidentification Approach can be adjusted to account for the anaphora phenomenon of (24) as well as (23) whereas Forward Depronominalization Approach cannot. In other words, since a pronoun cannot be an antecedent for another pronoun as Lakoff (1968) claims, the anaphoric relation of (24) should be analyzed as shown in (25).



The analysis of (25) is in accord with that of (23c) but not with that of (23b).

It has been claimed that the so-called backward pronominalization cannot proceed from the subordinate clause to the main clause. Thus, our NP Reidentification process should not be able to proceed from the main clause to the subordinate clause. Indeed, all the above-discussed cases of the so-called backward pronominalization or NP Reidentification are cases where the pronoun is found in the subordinate clause whereas the antecedent in the main clause. However, that is not always the case. Consider (26).

- (26) a. #*He* choked when *John* swallowed the bone.
 b. When *he* swallowed the bone, *John* choked.
 c. *He* choked—when *John* swallowed the bone, that is.

Assuming that *John* was mentioned in the preceding discourse, we can accept (26b, c), but not (26a), as showing the indicated coreferentiality. Note that in (26c) the backward pronominalization is from the subordinate clause to the main clause. We see here that something more than the from-main-to-subordinate-clause condition is involved in the so-called backward pronominalization. In fact, Bolinger (1977) argues that what is really involved in the so-called backward pronominalization is not the from-main-to-

subordinate-clause condition but a structural tightness condition. In other words, he argues, the structural relation between the subordinate and the main clauses is rather tight in (26a) but rather loose in (26b, c); therefore, the backward pronominalization or NP Reidentification is possible in the latter but not in the former. That is, in (26b, c) the structural relation between the subordinate and main clauses is loose enough to warrant reidentification of the once-pronominalized NP as a full NP later in the same sentence, whereas that is not the case in (26a). Note that a subordinate clause that follows its main clause is in a tight relation with the main clause as in (26a), whereas a subordinate clause that precedes its main clause is in a loose relation with the main clause as in (26b), and that a pause between the main and subordinate clauses would have the same loosening effect as preposing the subordinate clause, as we see in (26c).

A similar loosening effect by a pause can be observed in the following case of conjoined sentences.

- (27) a. #*He* lied to me and *John* was my friend.
 b. *He* lied to me, and *John* was my friend.

He and *John* can be coreferential in (27b), but not in (27a).

Certain conjunctive adverbials like *then* has the loosening effect, as we see in (28).

- (28) a. #*He* lost the money, and *John* found it again.
 b. *He* lost the money, and then *John* found it again.

He and *John* can be coreferential in (28b), but not in (28a).

A general statement, as opposed to a specific statement, can also have the loosening effect, as we see in (29).

- (29) a. #I caught *it* when *the virus* came along.
 b. I always catch *it* when *the virus* comes along.

The indicated coreferentiality obtains in (29b), but not in (29a).

Since this structural tightness condition on the NP Reidentification process can be motivated by the fact that an NP, which is more loosely related to the preceding coreferential pronoun, is more likely to be reidentified as a full NP, it also supports Hypothesis 3.

Our fourth hypothesis is (30).

- (30) *Hypothesis 4*: The marked anaphora can be functionally explained.

Since we have claimed that the unmarked anaphora is always forward (Hypothesis 3), the marked anaphora means backward anaphora. We have so far discussed cases of apparent backward anaphora, which are shown above to be really cases of forward discourse/contextual anaphora. There are, however, cases of genuine backward anaphora, and Hypothesis 4 claims that they are due to nonsyntactic, functional reasons. Consider (31).

- (31) a. Hey Gus! Tell *him* to come in if you see *Tom* out there, will you?
 b. ?Hey Gus! Tell *Tom* to come in if you see *him* out there, will you?

In offhand sentences like (31a), we cannot assume any prior discourse context and yet the pronoun *him* precedes its coreferential NP *Tom*, hence a case of genuine backward anaphora. Furthermore, (31b) is slightly awkward as an offhand sentence though it is a case of forward anaphora. Note that (31a, b) cannot be accounted for by any syntactic constraint like Langacker's (1969) primacy constraint on backward pronominalization since according to Langacker's constraint (31a) should be ungrammatical while (31b) is grammatical.

On the other hand, the following functional explanation for (31a, b) is quite plausible. An offhand sentence tends to present the major content of the message at the beginning and then the specific details toward the end. Thus, we can assume that in an offhand sentence NP's tend to be introduced as light as possible, e.g. as pronouns if possible, at the beginning and then reidentified with full details toward the end. Accordingly, as an offhand sentence (31a) is more natural than (31b)⁷.

Similar observations can be made on (32), a Korean version of (31).

- (32) a. Gus-ya! ϕ til-ə o-la-ko ha-ə, Tom-i
 Gus-VOC enter-COMP come-IMP-COMP tell-IMP Tom-NM
 pakk-e issi-myən.⁸
 outside-LOC is-if
 'Hey Gus! Tell him to come in if you see Tom out there, will you?'
 b. Gus-ya! Tom-lil til-ə o-la-ko ha-ə, ϕ pakk-e
 Gus-VOC Tom-AC enter-COMP come-IMP-COMP tell-IMP outside-LOC
 issi-myən.
 is-if

'Hey Gus! Tell Tom to come in if you see him out there, will you?'

(32a) is slightly more likely and natural as an offhand sentence than (32b), though both are grammatical.

On the other hand, in cases where offhand connotation is inappropriate, such backward anaphora is not allowed, as we see in (33).

- (33) a. If *the general* comes back, put *him* to death.

⁷ Linguists like Bolinger (1977) claim that even cases like (31a) are not cases of genuine backward pronominalization. According to Bolinger (1977), in (31a) the preceding *him* is not pronominalized by the following *Tom*, but is introduced on its own for independent reasons like offhand intent, and then later reidentified as a full NP, *Tom*. In this paper, however, we will assume that a pronoun preceding its coreferential NP is a case of genuine backward anaphora unless the speaker and hearer can agree in identifying its (possible) antecedent in the preceding discourse or context.

⁸ ϕ indicates the position of the null anaphor of *Tom*. The abbreviations are as follows: VOC = Vocative Marker; NM = Nominative Marker; AC = Accusative Marker; LOC = Locative Marker; COMP = Complementizer; DEC = Declarative Marker; IMP = Imperative Marker; ASP = Aspect Marker; PAST = Past Tense Marker.

- b. #If *he* comes back, put *the general* to death.

Such a sentence as (33a) can never be an offhand sentence since putting a general to death is an act of grave consequence unlike telling someone to come in; thus (33b), the offhand version of (33a), is inappropriate. This supports the functional explanation of the backward anaphora in cases like (31a) and (31b).

Another case of genuine backward anaphora is due to a literary style. Compare (34a) and (34b), which are due to Kuno (1975).

- (34) a. At 2 or 3 or 4 a.m., somewhere along in there, on August 25, 1966, *his* forty-eighth birthday, in fact, *Leonard Bernstein* woke up in the dark in a state of wild alarm.
 b. *Leonard Bernstein* woke up in the dark in a state of wild alarm at 2 or 3 or 4 a.m., somewhere along in there, on August 25, 1966, *his* forty-eighth birthday, in fact.

Both (34a) and (34b) can be an opening statement of a book, hence (34a) is a case of genuine backward anaphora. According to Kuno (1975), (34a) holds the reader in suspense and has a literary flavor whereas (34b) does not and gives the impression that the story is more documentary than literary.

Another case of genuine backward anaphora is due to the topical nature of subject, as we see in (35).

- (35) a. In *his* Memoirs, *Winston Churchill* tells us.....
 b. ?In *his* term paper, *Jerry Jones* tells us.....

Both (35a) and (35b) can be a context-independent, discourse-initial sentence, hence both are cases of genuine backward anaphora though some people regard (35b) as slightly awkward for the sense that *his* and *Jerry Jones* are coreferential.⁹ Note that it is most likely that the topics of the sentences (35a, b) are *Winston Churchill* and *Jerry Jones* respectively. If we assume that the topic of a sentence normally occurs at the subject position and that topic is never pronominalized intrasententially, or realized as a proform with its antecedent full NP within the same sentence as Kuno (1972) claims, then we can account for the backward anaphora in cases like (35a, b) as follows. That is, in (35a, b) the coreferential NP in the adverbial phrase, instead of the coreferential NP in the subject position, has to be in a proform even though the former precedes the latter, since the latter is the topic of the sentence.

⁹ Bolinger (1977) claims that (35a) is acceptable as a discourse-initial sentence but (35b) is not, and that the difference between them in acceptability is due to the fact that *Winston Churchill* is a well-known person whereas *Jerry Jones* is a nonentity. Many native speakers of English, however, accepts (35b), as well as (35a), as a discourse-initial sentence. Therefore, we cannot say that the acceptability of (35a) as a discourse-initial sentence, or a case of backward anaphora, is solely due to the fact that the subject *Winston Churchill* is a well-known person.

We have so far considered only three functional factors for backward anaphora: offhand usage, literary flavor and topical nature of subject. It can be assumed that there are other functional factors for backward anaphora, and that all the other cases of backward anaphora also will be accounted for by further functional, rather than syntactic, factors. Functional factors, unlike syntactic factors, will readily contribute to such higher level linguistic explanations as communicative functions of language or language use. Such functional studies on anaphora will also contribute to better understanding of syntactic aspects of anaphora by complementing the syntactic explanations for anaphora.

4. Lexical Heaviness of Anaphors

Our fifth hypothesis is (36).

- (36) *Hypothesis 5*: Lexically heavier anaphors are more constrained in coreference relation.

Hypothesis 5 is intuitively plausible in that lexically heavier anaphors would stand for antecedents of a narrower range than lexically less heavy ones. Consider (37a, b, c), which are due to Cole (1974).

- (37) a. # If *the centaurs* are feeling frisky, *some centaurs* chase unicorns.
 b. ?When *it* is cornered, *every zebra* will attack.
 c. Upon ϕ awakening, *some Martians* fail to brush *their* teeth.

(37a, b, c) are examples of backward anaphora with indefinite antecedents. (37a) is unacceptable for the sense in which the anaphor *the centaurs* refers to the antecedent *some centaurs*. In other words, definite descriptions like *the centaurs* cannot serve as backward anaphors of indefinite antecedents.

On the other hand, some, if not all, native speakers of English accept (37b) for the sense in which the anaphor *it* refers back to the indefinite antecedent *every zebra*. That is, for some speakers pronouns may serve as backward anaphors of indefinite antecedents. (37c) is acceptable for the sense in which the null anaphor refers to the indefinite antecedent *some Martians*. That is, the null anaphors can serve as backward anaphors of indefinite antecedents.

The different degrees of anaphoricity of a definite description, a pronoun and a null anaphor for an indefinite antecedent in backward anaphora are in accord with the different degrees of lexical heaviness of them. That is, lexically heavier anaphors have more restricted range of coreferentiality.

Similar phenomena can be observed in VP anaphora. Compare (38) and (39), which are due to Schachter (1977).

- (38) a. John expects to *get fired*, but I doubt he will ϕ .
 b. John expects to *get a raise*, but I doubt he will ϕ .

- c. John expects to *get a job*, but I doubt he will ϕ .
- (39) a. #John expects to *get fired*, but I doubt he'll *do it*.
 b. ?John expects to *get a raise*, but I doubt he'll *do it*.
 c. John expects to *get a job*, but I doubt he'll *do it*.

(38) and (39) show that the referential range of a *do it* VP anaphor is narrower than that of a null VP anaphor. That is, a null VP anaphor can stand for any kind of VP's, but a *do it* VP anaphor can stand for only volitional VP's like *get a job* but not for nonvolitional VP's like *get fired* or *get a raise*. This is what Hypothesis 5 predicts: the lexically heavier *do it* VP anaphor is more restricted in coreference relations than the lexically lighter or null anaphor ϕ . Take another example from Korean. Consider (40a, b, c).

- (40) a. *ki salam-i am-lil alh-nin kəs-lil John-i cal al-ko*
 the man-NM cancer-Ac have-COMP it-AC John-NM well know-COMP
iss-ta-ko Tom-i səngkakha-nin-ta
 is-DEC-COMP Tom-NM think-ASP-DEC
 'Tom thinks that John knows well that the man has a cancer.'
- b. *ki-ka am-lil alh-nin kəs-lil John-i cal al-ko*
 he-NM cancer-AC have-COMP it-AC John-NM well know-COMP
iss-ta-ko Tom-i səngkakha-nin-ta
 is-DEC-COMP Tom-NM think-ASP-DEC
 'Tom thinks that John knows well that he has a cancer.'
- c. ϕ *am-lil alh-nin kəs-lil John-i cal al-ko iss-ta-ko*
 cancer-AC have-COMP it-AC John-NM well know-COMP is-DEC-COMP
Tom-i səngkakha-nin-ta
 Tom-NM think-ASP-DEC
 'Tom thinks that John knows well that he has a cancer.'

The anaphor *ki salam* 'the man' in (40a) can never refer to *John* or *Tom* in the sentence but only to an antecedent identifiable in the discourse or context, whereas the null anaphor in (40c) can refer to not only the discourse/contextual antecedent but also *John* or *Tom* in the sentence. On the other hand, the anaphor *ki* 'he' in (40b) can normally refer to the discourse/contextual antecedent but also weakly to *John* or *Tom* in the sentence. It can be represented schematically as follows.

(41)

	Discourse Anaphora	Sentential Anaphora
<i>ki salam</i> 'the man'	O	X
<i>ki</i> 'he'	O	?
ϕ	O	O

This is exactly what Hypothesis 5 predicts: the lexically heavier anaphor *ki salam* is more restricted in coreference relations than the less heavier anaphor *ki*, which is in turn more restricted than the null anaphor ϕ .

Our last hypothesis is (42).

(42) *Hypothesis 6*: More constrained anaphors have a stronger coreference tie.

Both the anaphor *ki salam* in (40a) and the null anaphor ϕ in (40c) can refer to the discourse/contextual antecedent, as indicated in (41). But the coreference tie between the anaphor *ki salam* and the discourse/contextual antecedent in (40a) is felt stronger than that between the null anaphor and the discourse/contextual antecedent in (40c). This is exactly what Hypothesis 6 predicts: the lexically heavier and thereby more constrained anaphor *ki salam* has a stronger coreference tie than the lexically lighter or null anaphor ϕ .

For another piece of evidence for Hypothesis 6, consider (43).

- (43) a. John-i Tom-lil ki-iy cip-esə manna-əss-ta
 John-NM Tom-AC he-'s house-at meet-PAST-DEC
 'John met Tom at his house.'
- b. John-i Tom-lil caki-iy cip-esə manna-əss-ta
 John-NM Tom-AC self-'s house-at meet-PAST-DEC
 'John met Tom at his house.'

Note that the only formal difference between (43a) and (43b) is that *ki* 'he' is used in the possessive NP position in (43a) whereas *caki* 'self' is used in (43b). The pronominal anaphor *ki* 'he' can refer to *John* as well as *Tom* in (43a) whereas the heavier anaphor *caki* 'self' can refer to *John* but not *Tom* in (43b).¹⁰ On the other hand, the coreference tie between *John* and *caki* in (43b) is felt stronger than that between *John* and *ki* in (43a). But this is what Hypothesis 6 predicts: the lexically heavier, thereby more constrained anaphor *caki* has a stronger coreference tie than the lexically light and less constrained anaphor *ki*.¹¹

5. Summary and Conclusion

In this paper I have proposed the following six hypotheses for the functional approach to anaphora.

Hypothesis 1: Sentential anaphora and discourse/contextual anaphora, or linguistic anaphora and nonlinguistic anaphora, are one and the same phenomenon.

Hypothesis 2: Sentential anaphora is a special case of discourse/contextual anaphora.

¹⁰ This is what Hypothesis 5 predicts.

¹¹ The anaphor *ki* 'he' is clearly a pronoun, but the anaphor *caki* 'self' is of Chinese origin and more like a full noun. Anyway, it is clear that the latter is lexically heavier than the former.

Hypothesis 3: The unmarked anaphora is always forward.

Hypothesis 4: The marked anaphora can be functionally explained.

Hypothesis 5: Lexically heavier anaphors are more constrained in coreference relation.

Hypothesis 6: More constrained anaphors have a stronger coreference tie.

Since these hypotheses, especially hypotheses 1, 2, 3, and 4 are concerned with rather fundamental issues on anaphora, if they prove tenable there would be significant impacts on analyses of anaphora and the current grammatical theory in general. Even before a full proof of the hypotheses, it seems now clear that any theory or analysis of anaphora in a purely syntactic perspective can never be a complete characterization of anaphora in human language.

REFERENCES

- Bolinger, D. 1977. Pronouns and repeated nouns. Indiana University Linguistics Club.
- Cole, P. 1974. Indefiniteness and anaphoricity. *Language* 50. 665-674.
- Delisle, G.L. 1973. Discourse and backward pronominalization. Indiana University Linguistics Club.
- Hankamer, J. and I. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7. 391-428.
- Kuno, S. 1972. Functional sentence perspective. *Linguistic Inquiry* 3. 269-320.
- . 1975. Three perspectives in the functional approach to syntax. *Functionalism*. Chicago Linguistic Society. 276-336.
- Lakoff, G. 1968. Pronouns and reference. Indiana University Linguistics Club.
- Langacker, R. 1969. On pronominalization and the chain of command. In D. Reibel and S. Schane (eds.), *Modern Studies in English*. Englewood Cliffs, N. J.: Prentice-Hall. 160-86.
- Ross, J. 1969. On the cyclic nature of English pronominalization. In D. Reibel and S. Schane. 187-200.
- Schachter, P. 1977. Does she or doesn't she? *Linguistic Inquiry* 8. 763-767.
- Williams, E. S. 1977. On "Deep and Surface Anaphora". *Linguistic Inquiry* 8. 692-696.

Department of English

Seoul National University

San 56-1, Shinlim-dong, Kwanak-ku

Seoul 151, Korea

(Received 15 September 1978)