Toward a Feature-Movement Theory of Long-Distance Anaphora

Norvin Richards MIT April 1996

1. Introduction

A number of recent versions of binding theory have been based partly on notions like Shortest Move and Economy. According to binding theories of this kind, anaphors enter into some movement-like relation with their antecedents; Shortest Move is taken to account for locality restrictions on binding[1]. Furthermore, some notion of representational economy is invoked to account for Condition B; pronouns are taken to be less economical than anaphors, and are thus to be avoided whenever possible[2].

This paper will concentrate on two difficulties posed by long-distance anaphora for binding theories of this kind; these will be cases in which full obedience to Shortest Move and Economy appears to be optional. I will argue that these difficulties are only apparent and that binding theories can in fact be formulated along Minimalist lines.

2. Problems

2.1 The Shortest Move Problem

The use of Shortest Move to define binding domains leads us to expect that an anaphor ought to only have one possible

[1] Binding theories along these lines include Lebeaux (1983), Cole, Hermon, and Sung (1990), Huang and Tang (1991), Cole and Sung (1994), among many others.

[2] In particular, see much of the recent work of Burzio (1989, 1991, to appear) for proposals along these lines.

antecedent--the closest one. As the Icelandic example in 1 shows, long-distance anaphora raises problems for this expectation :

1. Jón heldur að Haraldur se að skrifa bókina sina

   Jon thinks that Harald is writing book self's

   'Jon$_i$ thinks that Harald$_j$ is writing his$_{i/j}$ book'

Here the anaphor *sina* 'self's' can refer to either of the c-commanding subjects--unexpectedly, in a Shortest Move-based theory. If *Haraldur* is a possible antecedent, then the shortest possible move for the anaphor is to *Haraldur*, and binding by *Jon* ought to be impossible.

In fact, the problem is even more acute than this, as can be seen in 2 (Icelandic, Höskuldur Thráinsson, p.c.):

2. Jón heldur að Haraldur se að skrifa bókina sina í

   Jon thinks that Harald is writing book self's in

   herberginu sinu

   room self's

   'Jon$_i$ thinks that Harald$_j$ is writing his$_j$ book in his$_i$ room'

2 contains two long-distance anaphors, each of which has a different antecedent. This argues against accounts of the ambiguity of 1 which invoke some kind of structural ambiguity, making *Haraldur* a possible antecedent in one structure but not in another. Here *Haraldur* is clearly a possible antecedent, as it is an <u>actual</u> antecedent; it binds the possessor of *bókina* 'book'. However, it is also transparent for binding; the possessor of *herberginu* 'room'

is bound by the higher subject *Jón*. This poses a serious problem for a Shortest Move-based theory of anaphora. If locality in anaphora is to be captured by some principle like "an anaphor must be bound by the closest possible antecedent", then why can the possessor of 'room' in 2 be bound by *Jón* when *Haraldur* is a closer possible antecedent?

2.2 The Economy Problem

Some binding theories have made use of some notion of Economy to deal with the facts in 3:

3. a. John$_i$ understands himself$_i$
   b. *John$_i$ understands him$_i$

According to binding theories of this kind, anaphors are taken to be more "economical", in some sense, than pronouns. The ill-formedness of 3.b is then attributed to the availability of a more structurally economical means of expressing the same proposition, namely 3.a.

The properties of long-distance anaphors raise a certain problem for this approach. Typically, long-distance anaphors are not in complementary distribution with pronouns (Icelandic, from Thráinsson 1991):

4. a. Jón sagði að ég hefði svikið sig
      John$_i$ said that I had betrayed self$_i$
   b. Jón sagði að ég hefði svikið hann
      John$_i$ said that I had betrayed him$_i$

If 3.b is ill-formed because 3.a is a more economical way of saying the same thing, we may wonder why 4.b is not ill-formed, given the availability of 4.a. On the assumption

that all anaphors are more economical than pronouns, 4.a ought to be preferable to 4.b.

Just as with the Shortest Move problem, we can see that structural ambiguity is probably not the right way of resolving the Economy problem. We might, for instance, be tempted to claim that subjunctive clauses in Icelandic are "optional binding domains", so that binding of the object of *svikið* 'betrayed' is possible in 4.a but impossible in 4.b. Such an approach will encounter difficulties with sentences like 5, as Rögnvaldsson (1986) and Maling (1984) both point out:

5. Jón taldi að María hefði sagt að ég hefði
   Jon_i believed that $Mary_j$ had said that I had
   skilað henni bókunum sínum
   returned $to\text{-}her_j$ books $self_i$'s

Here the lowest clause is transparent for long-distance binding of *sínum* 'self's' by *Jón*. We must apparently conclude that binding of anaphoric elements in the lowest clause is possible by members of the higher clauses. On the other hand, the lowest clause is opaque for binding as far as the pronoun *henni* 'her-DAT' is concerned. This pronoun can corefer with *María*, apparently obeying Economy, as the sentence is well-formed.

In short, a binding theory based on notions of Shortest Move and Economy seems to run into contradictions when it attempts to deal with long-distance anaphora. A rule like "anaphors must be bound by the closest antecedent" encounters

difficulties, since anaphors can seemingly skip possible (and even actual) antecedents. A rule like "use anaphors whenever possible" is also in trouble, as anaphors and pronouns are typically not in complementary distribution in this domain. A Minimalist approach to anaphora thus seems untenable.

In what follows I will attempt to develop a Minimalist approach to long-distance anaphora, drawing on Chomsky's (1995) notion of feature-movement, which will deal with these apparent problems.

3. Feature-Movement and Binding

Chomsky (1995) suggests that among the things that can move are features of various kinds. In a case like 6, for example, the feature [+wh] is claimed to move into CP in order to enter into a checking relation with a feature on C^{O}, forming a "feature-chain" headed by the wh-feature:

6. What did you say t?

In a language like Japanese, only the wh-feature need move to form a well-formed checking relation, and the wh-word itself therefore remains in situ. In English, on the other hand, a wh-feature cannot move by itself, presumably for morphological reasons. As a result, all the features associated with the wh-word *what* are "pied-piped" along with the wh-feature. This point of view makes feature-movement primary; whole words (or "categories") do move sometimes, for language-specific and idiosyncratic reasons, but this movement is always driven by feature-movement, and it is feature-movement with which syntax should largely concern

itself. Feature-movement is the kind of movement to which Shortest Move is taken to apply, for example.

The Minimalist approach to syntax centers largely on checking relations between specifiers of heads and features on the heads; syntactic movements are taken to be driven by a need to check off features of various kinds. In particular, in Chomsky 1995, feature-movement is performed in order to eliminate "uninterpretable" features; that is, features which give no instructions to either of the interfaces (PF or LF). The N-features on T(ense)$^{\circ}$, for example, are taken to be uninterpretable, and they therefore force raising of some nominal (namely, the subject) into the specifier of TP in order to eliminate the N-features before the interfaces are reached. The N-features on a nominal, on the other hand, are taken to be interpretable; no movement need be performed to get rid of them. The postulation of feature-chains raises interesting questions, in the Minimalist program, about what kinds of conditions exist on well-formed checking relations; what kind of relation does a feature-chain have to bear to a specifier for feature-checking to take place?

In this section I will suggest a partial answer to this question, examining particularly the behavior of the N-features on T$^{\circ}$ and on theta-assigners, which enter into relations with nominals. Crucially, my account will involve the claim that features can sometimes be base-generated "underspecified" for their value; in fact, as we will see, an Economy condition will require the base-generation of

underspecified features whenever possible. I want to claim that a feature-chain C entering into a checking relation with some element X is subject to the following restrictions:

7. a. Accuracy--C must be associated with a feature that accurately reflects the features on X

b. Completeness--C must be associated with a fully specified set of features.

"Association" is defined as follows:

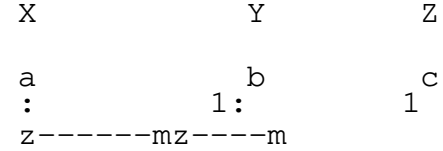
8. For any feature-chain C headed by the feature F and occupying the heads $\{H_1, H_2...\}$, for any feature G, C is associated with G iff

a. F=G

or b. some feature-chain occupying some element of $\{H_1, H_2...\}$ is associated with G

That is, a feature-chain C is associated with its own feature, as well as with the features of any feature-chains that occupy any of the heads occupied by C; moreover, association is transitive. Consider the feature-chains in 9:

9.

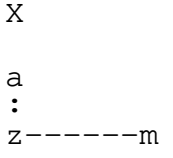


The features a, b, and c are base-generated on the heads X, Y, and Z, as shown. The feature c moves to the head Y, thus creating a feature-chain headed by c that connects Y and Z; similarly, the feature b moves to the head X. The feature-chains headed by a, b, and c are associated with the features a, b, and c respectively. Furthermore, a feature-chain is

associated with the features on any head to which it moves; thus, the feature-chain headed by c is also associated with the feature b, and the feature-chain headed by b is associated with a. Finally, association is transitive; because c is associated with b and b is associated with a, c is associated with a.

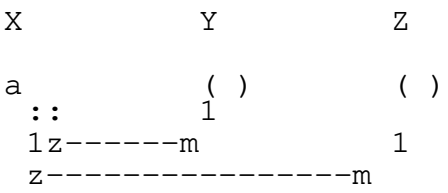
If a feature is base-generated underspecified, the requirement of Completeness will require that it move to another head where it can be associated with fully specified features. 10 gives an example of this:

10.

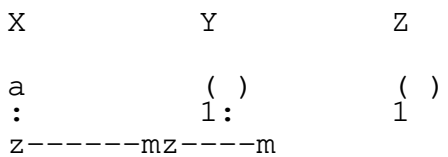


In 10, the underspecified feature on Y moves to X to be associated with the fully specified feature a on X. Consider what will happen in a structure with two underspecified features, as in 11:

11. a.



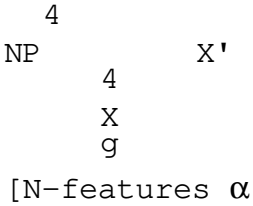
b.



In 11, there are two possible ways for both underspecified features to be associated with the fully specified feature on

X. Both features could move to X, as in 11.a. Alternatively, because association is transitive, the underspecified feature on Y could move to X, thus becoming associated with the fully specified feature on X, and the underspecified feature on Z could then move to Y, thus also becoming associated with a, by transitivity; this is shown in 11.b. If we follow Chomsky 1995 in assuming that feature-movement is subject to Shortest Move, then the choice between 11.a and 11.b will be entirely a question of which feature moves first. If the feature on Z moves first, then it will have to move to X, this being the closest head with a specified feature, and 11.a will be the result. If the feature on Y moves first, on the other hand, the feature on Z now need not move any further than Y to be associated with fully specified features; thus, Shortest Move prevents it from moving any further than Y, and we get the structure in 11.b.

The theory will make two claims about anaphors. First, all anaphors are taken to have underspecified N-features, so that Accuracy will require feature-chains checked by them to be associated with underspecified features. Second, the configuration in 12 is taken to be interpretable just in case the N-features α are those of an anaphor:

12.


Recall that in Chomsky's (1995) framework, features must be checked off and eliminated by the computational system because they are uninterpretable by the interfaces with the semantic and phonological components. N-features on theta-assigners and T^{O} are taken to be uninterpretable. I claim here that the interpretive component contains instructions assigning an interpretation to configurations like 12 involving some semantic relation between the NP in Spec XP and the anaphor that supplied the N-features; for reflexives, for example, this semantic relation will involve coreference. Just in case the N-features on a head are those of an anaphor, then, they are <u>not</u> uninterpretable, and therefore need not be checked off. The process of checking off N-features, then, will spare the N-features of anaphors, and the result of the derivation will be a tree bare of N-features except for those of anaphors.

Finally, I will assume that a principle of structural economy demands that underspecified features be used to express a given proposition whenever possible. Since anaphors have underspecified features and pronouns do not, this amounts to a requirement that anaphors be used whenever they can be.

I will be unable to develop a full account here of many of the phenomena for which binding theories are responsible; in particular, space limitations prevent me from discussing the typology of different kinds of anaphors found in different languages in any detail. In the theory developed

here, this typology is reflected in a typology of different types of N-features which may be underspecified in different anaphors. For purposes of this paper, I will concentrate on long-distance anaphors, which typically lack a feature that can only be supplied by T^{O}; since subjects are taken to move to Spec TP in this theory, long-distance anaphors are typically subject-oriented.

Before turning to an account of the facts of long-distance anaphora, however, let us see how the theory sketched thus far accounts for some basic binding facts. 13.a-b represent the core cases for Conditions A and B:

13. a. Bill_{i} thinks John_{j} criticized $\text{himself}_{*i/j}$

b. Bill_{i} thinks John_{j} criticized $\text{him}_{i/*j/k}$

In 13.a, Accuracy demands that the N-features on the verb be underspecified, so that feature-checking can take place with the underspecified features of the anaphor *himself*. Completeness then requires that these features move to T^{O}, where they can be associated with the fully-specified features on that head; this creates an N-feature-chain connecting T^{O} and the verb, which the interpretive component will assign an interpretation involving coreference between *John* and *himself*. Shortest Move demands that the N-features move to the <u>closest</u> T^{O}, which accounts for the inability of *Bill* to bind *himself*. 13.a and 13.b are alike in all relevant respects except for the fact that the verb *criticized* in 13.a bears underspecified N-features, while the corresponding verb in 13.b has fully specified features.

Economy thus prefers the structure in 13.a to 13.b. As a result, 13.b cannot have any reading which could be expressed more economically as 13.a; in other words, *him* cannot refer to *John*, although it can have any other referent.

Having seen how Economy and Shortest Move conspire to give us the salient properties of short-distance anaphora, let us finally turn our attention to long-distance anaphora, where these principles appear to be violated, as we have seen.

4. Long-Distance Anaphora

One peculiar property of the binding theory sketched above is the implicit distinction between the syntactic and semantic mechanisms involved in binding. Syntactically, binding involves feature-movement triggered by a need to set up well-formed checking relations. This feature-movement is purely syntactically motivated, and does not in itself entail any semantic consequences; the anaphor is not moving out of a need to seek an antecedent, for example. The semantic effects of binding are taken here to be the result of a separate interpretive mechanism that assigns readings of certain kinds to syntactic structures involving the movement of N-features of anaphors. In drawing this distinction between the syntax and the semantics of binding, I am leaving open the possibility of phenomena which exhibit the syntax of binding without its semantics; cases, that is, in which N-features move without triggering any of the interpretive

effects associated with anaphora. I want to claim that long-distance anaphora always involves such a case.

Consider again the Icelandic sentence in 4.a, repeated as 14:

14. Jón sagði að ég hefði svikið sig
John$_i$ said that I had betrayed self$_i$

In terms of the theory developed here, the portion of the structure of 14 which is relevant for binding theory consists of the N-features on three heads; the T^{O} bearing the N-features of *Jón*, the T^{O} bearing the N-features of *ég* 'I', and the verb *svikið* 'betrayed', which bears the underspecified N-features of the anaphor *sig*. The claim made here is that the lower T^{O} also bears underspecified features, despite not being in a checking relation with an anaphor:

15. Jón T^{O} **N(x)** sagði að ég T^{O} **N()** hefði svikið **N()** sig

:: 1

g z----------m 1

z----------m

John$_i$ said that I had betrayed self$_i$

Here some N-feature x is missing from both of the lower relevant heads, and must be supplied by feature-movement. First we move the N-features of the verb *svikið* 'betrayed' to the nearest head possessing them, which is the higher T^{O}; next, we move the N-features on the lower T^{O} to the higher T^{O}. All N-feature chains now satisfy Completeness, and the structure is well-formed. Both chains satisfy Shortest Move at the point in the derivation at which they are formed. The two chains are syntactically equivalent, both being motivated by a need to set up well-formed checking relations.

Semantically, of course, they are quite distinct; the chain connecting the two T^{O} heads is semantically vacuous, while the other chain triggers coreference between *Jón* and the anaphor.

The approach developed here thus crucially involves the claim that long-distance anaphora always involves a dependency between the binder of the anaphor and any subjects intervening between the anaphor and its binder. In 15, if the lower T^{O} had been base-generated with fully specified features, Shortest Move would have forced movement of the N-features of the anaphor to it, and long-distance binding could not have taken place. Some evidence that an approach of this kind is on the right track may be found in Chinese long-distance anaphora.

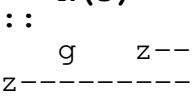
Consider what would happen if the N-feature missing on a long-distance anaphor were the feature Person. Any T^{O} heads intervening between the anaphor and the T^{O} binding the anaphor's features would have to be base-generated with underspecified Person features, so that Person-feature-movement could skip them. After the features of the anaphor had moved, the Person feature on each of the intervening T^{O} heads would move to the fully specified T^{O} head, in order to satisfy Completeness (Chinese, from Huang and Tang 1991):

16. Zhangsan $T^{O}_{\mathbf{N(3)}}$ renwei Lisi $T^{O}_{\mathbf{N()}}$ hai-le$_{\mathbf{N()}}$ ziji

Zhangsan$_i$ thinks Lisi hurt self$_i$

Of course, the Person-feature chains in 16 must satisfy Accuracy, as well as Completeness. In 16, this is achieved; the chain connecting the two T^{O} heads is associated with the 3rd person features on the higher T^{O}, and can therefore enter into a well-formed checking relation with the 3rd person lower subject, *Lisi*. If the lower subject were not 3rd person, however, Accuracy would not be satisfied, and the sentence would not be well-formed. This is in fact the case (Huang and Tang 1991):

17. *Zhangsan $T^{O}_{\mathbf{N(3)}}$ renwei wo $T^{O}_{\mathbf{N()}}$ hai-le$_{\mathbf{N()}}$ ziji



Zhangsan$_i$ thinks I hurt self$_i$

17 cannot have the reading shown; long-distance anaphora is impossible in this case. This illustrates the well-known "blocking effect" of Chinese; for long-distance anaphora to take place in Chinese, all the subjects intervening between the anaphor and its binder must agree with the binder in Person. This falls out straightforwardly from the system developed here, as we have seen, and the Chinese blocking effect thus provides some evidence for the claim that long-distance anaphora always involves dependencies between the long-distance binder and intervening subjects.

The presence of a blocking effect for Person in Chinese, but not in Icelandic, indicates the fact that long-distance anaphors are anaphoric for Person in Chinese, but not in Icelandic, according to this system. Supporting evidence for this may be found in the fact that Chinese *ziji* can be bound

by an antecedent of any person, while Icelandic *sig* must have a third-person antecedent:

18. (Icelandic, adapted from Thráinsson 1975):

a. *Ég hata sig
I$_i$ hate self$_i$

b. Jón hatar sig
John$_i$ hates self$_i$

19. (Chinese)

a. Wo hai-le ziji
I$_i$ hurt-ASP self$_i$

b. Zhangsan hai-le ziji
Zhangsan$_i$ hurt-ASP self$_i$

This is consistent with the claim that Icelandic *sig* enters the derivation bearing a third-person feature, while Chinese *ziji* lacks a Person feature.

5. The Shortest Move and Economy Problems

Now we are in a position to take another look at the problems raised for Shortest Move and Economy by long-distance anaphora. Recall that the problems had to do with the seemingly ambiguous status of intervening subjects for long-distance anaphors; intervening subjects seemed to be able to simultaneously act as potential binders and also be transparent for binding. The Shortest Move problem had to do with the well-formedness of sentences like 2, repeated as 20:

20. Jón heldur að Haraldur se að skrifa bókina sina í
Jon thinks that Harald is writing book self's in
herberginu sinu
room self's
'Jon$_i$ thinks that Harald$_j$ is writing his$_j$ book in his$_i$ room'

20 seems to raise problems for any binding theory that makes crucial use of some notion of Shortest Move to ensure that anaphors must be bound by the closest possible binder. Here *Haraldur* is clearly a possible binder, since it is in fact an actual binder of the possessor of *bókina* 'book', yet the possessor of *herberginu* 'room' can be bound by the higher subject *Jón*, skipping *Haraldur*. *Haraldur* thus appears to have a contradictory status with regard to Shortest Move; it is and is not a potential binder.

The Economy problem is similar in nature, and has to do with the well-formedness of sentences like 5, repeated as 21:

21. Jón taldi að María hefði sagt að ég hefði
Jon$_i$ believed that Mary$_j$ had said that I had
skilað henni bókunum sínum
returned to-her$_j$ books self$_i$'s

Here we have a problem for any binding theory that uses Economy to predict a preference for anaphors over pronouns whenever anaphors can be used. Clearly, long-distance binding into the lowest clause in 21 is possible, since there is a bound long-distance anaphor in that clause. However, there is also a pronoun in that clause which appears to be

bound within the domain of long-distance anaphora; that is, bound within a domain in which anaphora is possible, contrary to an Economy-based theory. On the assumption that what blocks movement to a particular binder is the availability of a closer potential binder, this can also be seen as reflecting an ambiguous status for the lowest subject in 21: *ég* 'I' both is and is not a potential binder, and the lowest clause thus both is and is not opaque for long-distance binding.

In the analysis of long-distance anaphora developed here, both of these problems are straightforwardly solved. Intervening subjects in cases of long-distance anaphora enter into dependencies with higher subjects. Thus, they begin the derivation with underspecified features, and are unable to bind; by the end of the derivation, their features have moved to a head with fully specified features, which renders them capable of binding. Their status as potential binders thus changes in the course of the derivation, which accounts, in this theory, for their ambiguous status.

The derivation of a sentence like 20 (repeated as 22.a) is given in 22.b-d[3]:

[3] Here I assume arbitrarily that a possessor is assigned its theta-role by the possessed noun. Nothing hinges on this, of course.

22. a. Jón heldur að Haraldur se að skrifa bókina sina

Jon thinks that Harald is writing book self's

í herberginu sinu

in room self's

'Jon$_i$ thinks that Harald$_j$ is writing his$_j$ book in his$_i$ room'

b. Jón $T^{O}_{\mathbf{N(x)}}$ heldur að Haraldur $T^{O}_{\mathbf{N()}}$ se að skrifa

:

1 bókina$_{\mathbf{N()}}$ sina í herberginu$_{\mathbf{N()}}$ sinu

z----------m

c. Jón $T^{O}_{\mathbf{N(x)}}$ heldur að Haraldur $T^{O}_{\mathbf{N()}}$ se að skrifa

:: 1

1z----------m

1

1 bókina$_{\mathbf{N()}}$ sina í herberginu$_{\mathbf{N()}}$ sinu

z----------m

d. Jón $T^{O}_{\mathbf{N(x)}}$ heldur að Haraldur $T^{O}_{\mathbf{N()}}$ se að skrifa

:: 1:

1z----------m1

1 a----------m

1 bókina$_{\mathbf{N()}}$ sina í herberginu$_{\mathbf{N()}}$ sinu

z----------m

The sentence is base-generated as in 22.b, with some N-feature x specified only on the highest T^{O}. In the first step, the N-features of the possessor of *herberginu* 'room' move to this T^{O}, obeying Shortest Move (since no closer head has the feature in question), creating a feature-chain that satisfies Completeness and can enter into a well-formed checking relation with the anaphor *sinu*, and triggering coreference between this anaphor and the highest subject,

Jón. In the second step, the features on the lower T^{O} raise to the higher T^{O}, again in order to create a well-formed checking relation. Finally, the features of the possessor of *bókina* 'book' must raise to become associated with fully specified features, thus satisfying Completeness. Because association is transitive, the closest head that will serve as a landing site for that purpose is the lower T^{O}; thus, Shortest Move forces the features of this anaphor to raise to that T^{O}, as shown in 22.d. Shortest Move is therefore satisfied at all points in the derivation, and the seemingly contradictory nature of the intervening subject is a matter of its status having changed in the course of the derivation.

The Economy problem is solved in a similar way. The derivation of a sentence like 21 (repeated as 23.a) is as in 23.b-d:

23. a. Jón taldi að María hefði sagt að ég hefði
Jon$_i$ believed that Mary$_j$ had said that I had

skilað henni bókunum sínum
returned to-her$_j$ books self$_i$'s

b. Jón $T^{O}_{\mathbf{N(x)}}$...María $T^{O}_{\mathbf{N()}}$...ég $T^{O}_{\mathbf{N()}}$...

skilað$_{\mathbf{N(x)}}$ henni bókunum$_{\mathbf{N()}}$ sínum

c. Jón $T^{O}_{\mathbf{N(x)}}$...María $T^{O}_{\mathbf{N()}}$...ég $T^{O}_{\mathbf{N()}}$...

skilað$_{\mathbf{N(x)}}$ henni bókunum$_{\mathbf{N()}}$ sínum

d. Jón $T^{O}_{\mathbf{N(x)}}$...María $T^{O}_{\mathbf{N()}}$...ég $T^{O}_{\mathbf{N()}}$...

skilað$_{\mathbf{N(x)}}$ henni bókunum$_{\mathbf{N()}}$ sínum

The derivation of 23.a is comparatively simple. As before, the anaphor's features move to the highest T^{O}, obeying Shortest Move and creating a binding relation between the highest subject and the anaphor. Next both of the lower T^{O} heads are bound by the highest one, creating well-formed feature-checking relations and allowing feature-checking to take place. There is a point in the course of the derivation at which the fully specified features of the pronoun are in fact in violation of Economy; after the step depicted in 23.c, where the T^{O} head associated with *María* has satisfied Completeness but the T^{O} head immediately below it has not.

At this point, if the pronoun *henni* were an anaphor, its features could move to the T^{o} associated with *María*; thus, this proposition is not being expressed in the most economical way possible. By the end of the derivation, however, Economy is satisfied again. If the pronoun *henni* were an anaphor, it could not be bound by *María*. Its features would only need to move to the closest T^{o}, the one associated with *ég* 'I', to satisfy well-formedness conditions, and would therefore be prevented by Shortest Move from moving any further. If we assume that Economy applies at the interface, then, after all binding relations have been established, the Economy problem goes away.

6. Shortest Moving languages

Before I conclude, I will briefly discuss a class of languages in which the Shortest Move problem noted above does not seem to arise. Recall that in languages like Icelandic, a long-distance anaphor can apparently be bound past a closer possible antecedent (22.a, repeated as 24):

24. Jón heldur að Haraldur se að skrifa bókina sina
Jon thinks that Harald is writing book self's
í herberginu sinu
in room self's
'Jon$_i$ thinks that Harald$_j$ is writing his$_j$ book in his$_i$ room'

Here the possessor of *herberginu* 'room' is bound by *Jón*, even though *Haraldur* is clearly a closer potential antecedent for long-distance anaphors (since it is, in fact, an actual

antecedent for a long-distance anaphor, namely the possessor of *bókina* 'book'). I have been using Icelandic examples to illustrate this problem. Similar facts are found in German, Danish, and Norwegian:

25. (German, Uli Sauerland, p.c.)

Hans ließ Maria ein Bild von sich
Hans let Maria a picture of self
an einen Freund von sich schicken
to a friend of self send
'Hans$_i$ let Maria$_j$ send a picture of herself$_j$ to his$_i$ friend'

26. (Danish, Sten Vikner, p.c.)

?Peter bad Michael snakke med sin kone
Peter asked Michael talk with self's wife
om sine problemer
about self's problems
'Peter$_i$ asked Michael$_j$ to talk with his$_i$ wife about his$_j$ problems'

27. (Norwegian, Arild Hestvik, p.c.)

?John lot Chomsky skrive anbefalingen sin
John let Chomsky write recommendation self's
på sitt kontor
in self's office
'John$_i$ let Chomsky$_j$ write his$_i$ recommendation in his$_j$ office'

On the other hand, there are some languages in which readings of this kind are strongly disfavored. Howard and Howard

(1976) note, for example, that multiple instances of *zibun* in Japanese must typically corefer if they share a set of possible binders. Thus, 28.a can have the readings in 28.b-c, but not those in 28.d-e:

28. a. Taroo-wa [Hanako-ga zibun-no heya-de
Taroo TOP Hanako NOM self GEN room in
zibun-no sigoto-o site-ita to] itta
self GEN work ACC do ing that said

b. Taro$_i$ said that Hanako$_j$ was doing his$_i$ work in his$_i$ room

c. Taro$_i$ said that Hanako$_j$ was doing her$_j$ work in her$_j$ room

d. *Taro$_i$ said that Hanako$_j$ was doing his$_i$ work in her$_j$ room

e. *Taro$_i$ said that Hanako$_j$ was doing her$_j$ work in his$_i$ room

Similar facts in Korean are discussed by Fiengo and Kim (1980)[4] (Yoo-Kyung Baek, p.c.):

[4] Fiengo and Kim (1980) present a very interesting set of facts, which I cannot do justice to here, partly because my informants disagree with several of their crucial judgments. I will have to leave this problem to further research.

29. a. Chelswu-ka [Sunhi-ka caki pang-ese caki il -ul
Chelswu NOM Sunhi Nom self room in self work ACC
ha -ko -iss-ta -ko] sayngkakha-n -ta
do-Prog-Be -Dec-Comp think -Prs-Dec

b. Chelswu$_i$ said that Sunhi$_j$ was doing his$_i$ work
in his$_i$ room

c. Chelswu$_i$ said that Sunhi$_j$ was doing her$_j$ work
in her$_j$ room

d. *Chelswu$_i$ said that Sunhi$_j$ was doing his$_i$ work
in her$_j$ room

e. *Chelswu$_i$ said that Sunhi$_j$ was doing her$_j$ work
in his$_i$ room

Finally, a similar restriction is found in Chinese (Hooi-Ling Soh, p.c.):

30. a. Xiao Ming xiang Da Hua zai ziji de fangjian zuo
Xiao Ming think Da Hua at self 's room do
ziji de gongke
self 's homework

b. Xiao Ming$_i$ said that Da Hua$_j$ was doing
his$_i$ homework in his$_i$ room

c. Xiao Ming$_i$ said that Da Hua$_j$ was doing
his$_j$ homework in his$_j$ room

d. *Xiao Ming$_i$ said that Da Hua$_j$ was doing
his$_i$ homework in his$_j$ room

e. *Xiao Ming$_i$ said that Da Hua$_j$ was doing
his$_j$ homework in his$_i$ room

Thus, there appears to be a class of languages in which a single binder must be chosen as the binder for all long-distance anaphors in the sentence; a class of languages, that is, in which the Shortest Move problem does not arise. How can we characterize this class in terms of the theory developed here?

First of all, it is probably worth noting that the effect described above can be overruled by semantic factors, at least for some speakers (Korean, Yoo-Kyung Baek, p.c.):

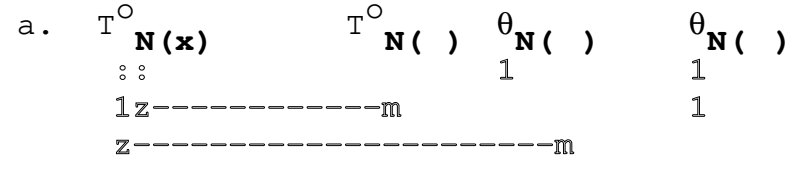
31. John-un [Bill-i [caki-ka caki-pota te
John TOP Bill NOM self NOM self than more
ttokttokha-ta-ko] sayngkakha-n -ta -ko] malha-ess-ta
smart Dec Comp think Prs Dec Comp say Past Dec
'John$_i$ thinks Bill$_j$ said self$_j$ is smarter than self$_i$'

Here the two instances of the long-distance anaphor *caki* can fail to corefer, presumably because the reading where they corefer woyuld be semantically anomalous (since people cannot be smarter than themselves). Thus, we will still need to allow the derivation that gave rise to the Shortest Move problem to be available, albeit disfavored, in these languages.

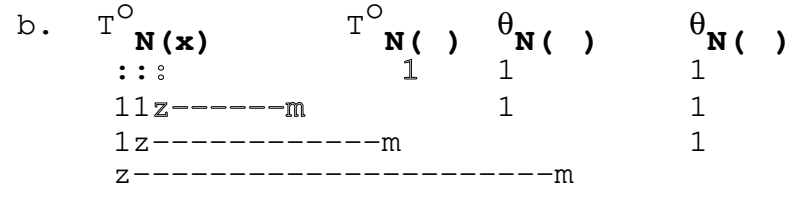
Let us consider the class of derivation which is disfavored in these languages. The sentences above involve two long-distance anaphors, each of which has a choice between two possible T^{o} heads as binders. Such a configuration allows four possible derivations in the theory we have developed here (and, as we have seen, all four

derivations are possible in some languages). The two which are permissible in languages like Chinese are described below. In the first, the features of both anaphors raise from the theta-assigners where they are base-generated to be bound by the higher T^{O}, after which the lower T^{O} is bound by the higher T^{O}. In the second, the higher T^{O} binds the lower T^{O} first, after which both of the anaphors have their N-features bound by the lower T^{O}:

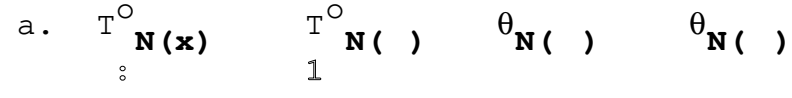
32. a.



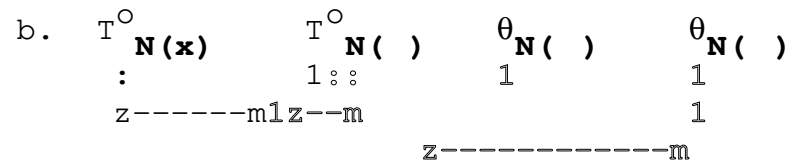
b.



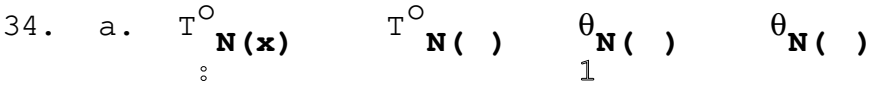
33. a.

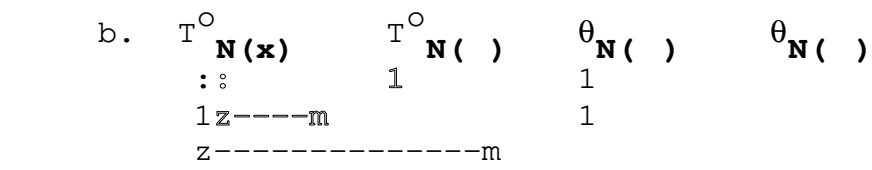


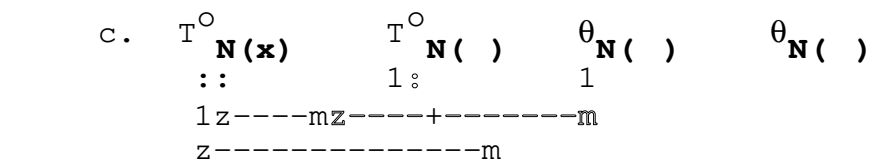
b.

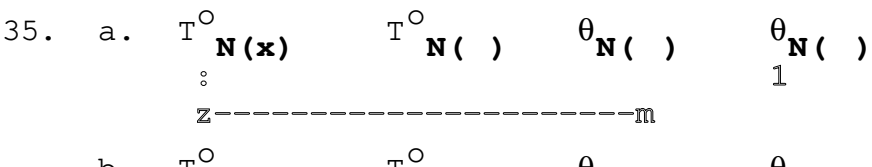


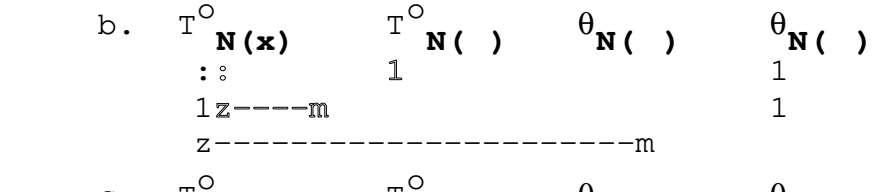
The other two derivations proceed as follows. First, one anaphor is bound by the higher T^{O}. Next, the lower T^{O} is bound by the higher T^{O}. Finally, the second anaphor is bound by the lower T^{O}. This is the kind of derivation which is allowed in some languages, including Icelandic, but disfavored in others, including Chinese:



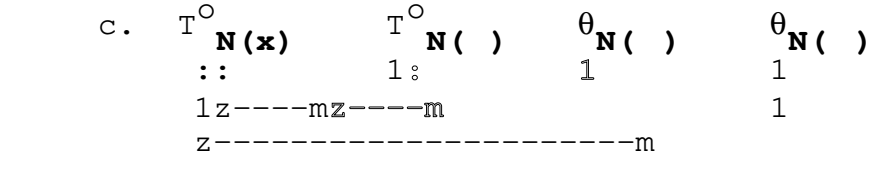








```
34.  a.  T°N(x)      T°N( )     θN( )      θN( )
         :                       1
         z--------------m

     b.  T°N(x)      T°N( )     θN( )      θN( )
         ::          1          1
         1z----m                1
         z--------------m

     c.  T°N(x)      T°N( )     θN( )      θN( )
         ::          1:         1
         1z----mz----+-------m
         z--------------m

35.  a.  T°N(x)      T°N( )     θN( )      θN( )
         :                                 1
         z---------------------m

     b.  T°N(x)      T°N( )     θN( )      θN( )
         ::          1                     1
         1z----m                           1
         z---------------------m

     c.  T°N(x)      T°N( )     θN( )      θN( )
         ::          1:         1          1
         1z----mz----m                     1
         z---------------------m
```



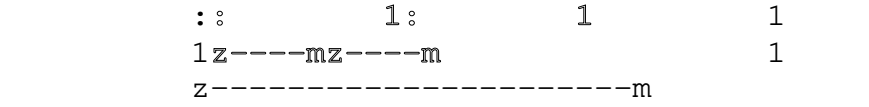
Descriptively speaking, then, languages like Chinese seem to prefer derivations in which binding of anaphors and binding of T^{o} heads occur in separate blocks; that is, derivations in which all anaphors are bound, followed by binding of all T^{o} heads, or in which all T^{o} heads are bound, followed by binding of all anaphors. Derivations in which these two kinds of binding are intermixed, as in 34-35, appear to be disfavored in this kind of language.

Obviously, this is the kind of restriction which we would want to explain by appeal to some other principle, and I have no suggestions to offer along those lines at this

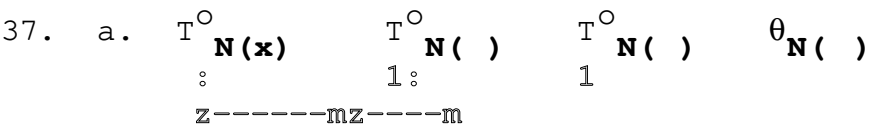
point. As a description of the phenomenon, however, this notion makes a further prediction in the context of this theory, which is borne out, as it happens. Consider a case in which a single anaphor has a choice of more than two T^{O} heads as possible binders. The theory developed above allows us to describe derivations in which the long-distance anaphor is bound by any of the c-commanding subjects. For example, the features of the anaphor could move first, followed by binding of all the T^{O} heads; this would result in a reading in which the anaphor was bound by the highest subject, as in 36. Alternatively, all the T^{O} heads could be bound first, followed by binding of the anaphor's features; this would give us the reading in which the anaphor was bound by the closest subject, as in 37. Finally, one T^{O} head could be bound by the highest T^{O} head, followed by binding of the anaphor, followed by binding of the lowest T^{O} head. This would result in a reading in which the anaphor is bound by the intermediate subject, as in 38:

36. a. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$

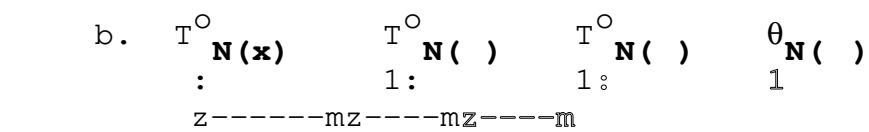
b. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



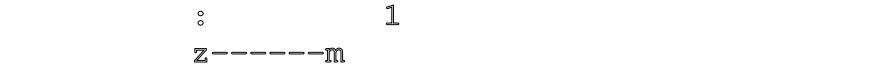
37. a. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



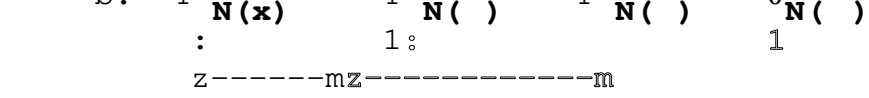
b. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



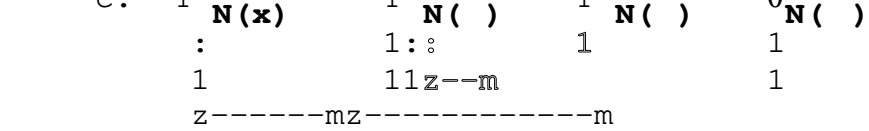
38. a. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



b. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



c. $T^{O}_{\mathbf{N(x)}}$ $T^{O}_{\mathbf{N(\ )}}$ $T^{O}_{\mathbf{N(\ )}}$ $\theta_{\mathbf{N(\ )}}$



If the descriptive claim made above is correct, Chinese should prefer derivations like those in 36 and 37, in which binding of the T^{O} heads occurs in a single block, to that in 38, where anaphor-feature-binding and T^{O}-binding are interspersed. That is, Chinese should prefer long-distance anaphors to be bound either by the closest possible antecedent, or by the furthest one. On the other hand, no such effect should appear in Icelandic. This appears to be the case. Battistella and Xu (1990), among others, have

noted that in a Chinese sentence like 39, binding by the intermediate subject is strongly dispreferred:

39. Zhangsan yiwei Lisi zhidao Wangwu bu xiangxin ziji
Zhangsan think Lisi know Wangwu not like self
'Zhangsan$_i$ thinks Lisi$_j$ knows Wangwu$_k$ doesn't like self$_{i/*?j/k}$'

Huang and Tang (1991) note that this effect can be countered by pragmatic factors, and it appears to pattern in that regard with the requirement that multiple long-distance anaphors share a binder, as noted above. As predicted, the effect does not appear in Icelandic[5] (Anderson 1986):

40. Jón segir að Maria viti að Haraldur vilji
Jon says that Maria knows-SUBJ that Haraldur wants-SUBJ
að Bill meiði sig
that Bill hurts-SUBJ self
'Jon$_i$ says that Maria$_j$ knows that Harald$_k$ wants Bill$_l$ to hurt self$_{i/j/k/l}$'

Thus, this theory allows a single descriptive statement to account for a range of derivations which are available in Icelandic but not in Chinese. As I remarked before, of course, this descriptive statement is somewhat unsatisfying, and we would like to develop a theory from which it will

[5] Rögnvaldsson (1986) notes cases in which binding by intermediate subjects is degraded in Icelandic, but the phenomena do not appear to be comparable to Chinese; the main factor seems to be the factivity of the verbs in the sentence. The contrast between Chinese and Icelandic thus seems to remain clear, although I will have nothing to say here about his observations.

follow naturally, but I will have to leave this for further research.

7. Conclusion

Let me close, then, with a quick overview of the solution proposed here to the problem of optionality in long-distance anaphora. It has been noted that principles which pick the "best" derivation, such as Economy and Shortest Move, encounter difficulties when dealing with cases in which several different derivations are apparently possible. I have sketched two particular difficulties encountered by Minimalist theories of anaphora in the realm of long-distance anaphora. I have then gone on to develop a theory of anaphora which is based on Economy and Shortest Move, but which deals successfully with these problems. The basic claim is that the syntactic mechanism which underlies the semantic relation between an anaphor and its antecedent can also take place between semantically unrelated elements; anaphoric binding is a particular application of a general operation of feature-binding. In particular, there are cases in which T^{O} is bound despite not being in a checking relation with an anaphor. In such cases, the status of T^{O} with respect to binding theory changes in the course of the derivation. It begins the derivation with deficient N-features, unable to bind heads lacking those N-features that it lacks itself, and thus transparent for movement of those N-features. By the end of the derivation, the N-features of the T^{O} are associated with fully specified N-features, and

are thus rendered able to bind. The optionality associated with long-distance anaphora, in this theory, thus has to do with an optionality in the order of operations in a derivation. If anaphors are bound before an intervening T^{O} is bound, then they can and must be bound by a higher T^{O}. If the intervening T^{O} is bound first, then anaphors are forced by Shortest Move to move to the intervening T^{O}. This theory thus crucially rejects any principles which rule out this kind of optionality in order of operations; for example, a strict version of the Strict Cycle condition, which allows no countercyclic movements at all, would be incompatible with this proposal. We must also crucially assume that principles like Shortest Move and Economy apply at particular stages in the derivation, rather than globally. In the theory developed here, for example, Shortest Move applies at the point in the derivation at which movement is performed, and later operations can change the representation in such a way that a move which was performed at the beginning of the derivation would be impossible at the end. This is crucial to the account developed above of the Shortest Move problem; anaphors can skip potential and even actual binders, as long as those binders were underspecified and thus unable to bind at the point in the derivation at which movement of anaphoric N-features took place.

Bibliography

Anderson, Stephen. 1986. The typology of anaphoric dependencies: Icelandic (and other) reflexives. Topics in Scandinavian syntax, ed. by Lars Helland and Kirsti Koch Christensen. Dordrecht: D. Reidel.

Battistella, Edwin, and Yonghui Xu. 1990. Remarks on the reflexive in Chinese. Linguistics 28.205-240.

Burzio, Luigi. 1989. On the non-existence of disjoint reference principles. Rivista di grammatica generativa 14.3-27.

Burzio, Luigi. 1991. The morphological basis of anaphora. Journal of Linguistics 27.81-105.

Burzio, Luigi. to appear. The role of the antecedent in anaphoric relations. to appear in Current issues in comparative grammar, ed. by Robert Freidin. Dordrecht: Kluwer.

Chomsky, Noam. 1995. Categories and transformations. ms., MIT.

Cole, Peter, Gabriella Hermon, and Li-May Sung. 1990. Principles and parameters of long-distance reflexives. Linguistic Inquiry 21.1-22.

Cole, Peter, and Li-May Sung. 1994. Head-movement and long-distance reflexives. Linguistic Inquiry 25.355- 406.

Fiengo, Robert, and Heasun Kim. 1980. Binding and control in Korean: structural restrictions on anaphora in a non-configurational language. Journal of Linguistic Research 1.59-73.

Howard, I., and A. Niekawa Howard. 1976. Semantics of the Japanese passive. Syntax and semantics 5: Japanese generative grammar, ed. by Masatoshi Shibatani. New York: Academic Press.

Huang, C.-T. James, and C.-C. Jane Tang. 1991. The local nature of the long-distance reflexive in Chinese. Long-distance anaphora, ed. by Jan Koster and Eric Reuland. Cambridge: Cambridge University Press.

Lebeaux, David. 1983. A distributional difference between reciprocals and reflexives. Linguistic Inquiry 14.723-730.

Maling, Joan. 1984. Non-clause-bounded reflexives in modern Icelandic. Linguistics and Philosophy 7.211-241.

Richards, Norvin. 1994. Head-binding. Proceedings of WCCFL 12, ed. by Erin Duncan, Donca Farkas, and Philip Spaelti. Stanford: Center for the Study of Language and Information.

Richards, Norvin. 1995. Towards a theory of head-binding. ms., MIT.

Rögnvaldsson, Eiríkur. 1986. Some comments on reflexivization in Icelandic. Topics in Scandinavian syntax, ed. by Lars Helland and Kirsti Koch Christensen. Dordrecht: D. Reidel.

Thráinsson, Höskuldur. 1975. Some arguments against the interpretive theory of pronouns and reflexives. ms., Harvard University.
Thráinsson, Höskuldur. 1991. Long-distance reflexives and the typology of NPs. Long-distance anaphora, ed. by Jan Koster and Eric Reuland. Cambridge: Cambridge University Press.