**The syntax of anaphora**

By Kenneth Safir


Reviewed by Erwin R. Komen

Radboud University Nijmegen

---

**Introduction**

This book is about anaphora, which is defined as “the relation established by grammatical forms that refer to previously established referents in discourse”. Anaphora in the broader sense plays an important role on the level of a whole text (discourse). This book, however, limits itself to describe a theory on anaphora within the domain of the sentence. Most of the book is about anaphora that requires a local antecedent, such as reflexives and reciprocals. Safir’s companion volume, entitled The syntax of (in)dependence, is said to go beyond the sentence level.

A further restriction to The syntax of anaphora is in the word categories treated. Most attention goes to personal, possessive and clitic pronouns as well as to quantifier expressions. Other types of Noun Phrases are only touched upon—I have not, for instance, found comments on the anaphoric behavior of demonstratives.

The theoretical framework of this book is generative grammar. The rules of anaphoric behavior are couched within the framework of the Minimalist Program (MP). Some knowledge of generative grammar and basic knowledge of the MP is needed, especially if you want to read chapters 5–7, whose topics are, I think, more related to theory internal matters.

The crucial idea defended in The syntax of anaphora is that only a theory that takes into account competition between representations can successfully account for the anaphora data. It is this point of view that makes the book worthwhile for a much wider audience than generative grammarians.

Chapter 1 introduces the research program of the book, which, basically, is to find the rules for the distribution of anaphoric forms and interpretations at the sentence level. The binding theory introduced in 1981 (Chomsky) explains locality effects for anaphora as local phenomena that are subject to c-command. Differences within and between languages are lexically determined. While rejecting binding theory, Safir defends this general lexicalist approach. Safir’s approach starts with the form to interpretation principle (FTIP), which is explained in more detail in chapter 3. The FTIP is a procedure that lays the relation between dependent forms and an identity interpretation. It compares different forms on a dependency scale, and explains distributional differences by a competition within the FTIP.
2. The interpretation of anaphoric relations from syntactic form

This chapter starts by distinguishing between dependent identity and coreference. When the interpretive content of form A depends on the interpretation of form B, there is dependent reference, as for example ‘his mother’ in (1) depends on ‘he’. When the referents of A and B necessarily are the same, then there is dependent identity, as illustrated in (1) by the relation between ‘his’ and ‘he’. When A and B have the same referents, but this is not out of necessity, then there is coreference. This is the case between ‘he’ and ‘John’ in (2) for instance.

1. He loves his mother.
2. John was introduced to the new secretary yesterday, but he already knew her.

The difference between dependent identity and coreference is a crucial one, since it is only dependency relations that formal grammar regulates directly, while coreference relations are regulated indirectly.

Chapter two convincingly argues that indices should be abandoned, since they represent symmetric coreference, whereas from a theoretical point of view only asymmetric dependent identity can be defended. The relation between form and interpretation is an asymmetric one, as explained in the next chapter.

3. Competition and complementarity

The form to interpretation principle (FTIP) introduced in chapter 3 leads into the competition principles. The algorithm works as follows:

- For each antecedent:
  - Consider all potentially dependent forms in the syntactic context.
  - Determine the form that is most dependent.
  - This form receives the interpretation of “dependent identity” (for the selected antecedent).

A natural consequence of comparing forms on a dependency scale is the complementary distribution between more and less dependent forms. The decreasing dependency scale reflexive anaphor > pronoun > r-expression, for instance, combined with the algorithm above, is capable of deriving the effects that have been formulated as Principle A, B and C within generative grammar. Particular languages have slightly different dependency scales, as the available forms differ from language to language. Germanic languages, for instance, have SIG and SELF forms on the scale.

Chapter 3 adds several principles to the FTIP that are worth noting. The coargument dependency constraint (CDC) deals with limitations on a distributed interpretation. The local antecedent licensing (abbreviated as LAL) restricts anaphors to a local domain. The remainder of chapter 3 tackles situations where complementary distribution seems to fail, explaining why there are independent reasons to account for these supposed failures.
Chapter 4 argues that not all dependent identity interpretations are alike, in particular when considering arguments of the same predicate (e.g. a subject and a reflexive). This is partly explained by the fact that the form with the dependent identity reading may have a different position on the dependency scale. The following interpretations are mentioned: codependent covaluation, dependent identity, indistinctness, guise and proxy.

Not all of these interpretations seem to be strictly dependent identity, witness the relation between ‘he’ and ‘Grisham’ in (3). The pronoun ‘he’ refers to a book written by Grisham, which is translated into Swahili, whereas ‘Grisham’ refers to the author, John Grisham.

3. Grisham claims that he is even more suspenseful in Swahili.

Chapter 5 is quite technical and attacks a problem that seems to be closely related to the theoretical framework—that of generative grammar. The chapter tackles another aspect of anaphors (in their narrow sense), namely that of the size of their domain. Safir argues that their local domain should be restricted. He notes that there are discrepancies with observed dat, and proposes to explain these in a way that is quite fashionable for the MP. Wherever they show up in unexpected places, they must have moved in the covert (invisible) part of grammar. For their interpretation at LF (Logical Form), they are in one place, but as far as their realization (that is: PF) is concerned, they are found in another place. The different behaviors of different anaphors are then attributed to lexical differences.

While these explanations certainly suffice within the MP framework, I personally find it hard to see what they contribute to our understanding of the processes involved.

Chapter 6 deals with the relation between the morphological makeup of anaphors (again, in the narrow sense) and their distribution. Safir moves away from a rigid internalist approach, which says that there should be a direct relation between the internal makeup of a form and its distribution. Instead, he argues for a two-way process. The internal properties (the morphological makeup) of different forms (such as strong pronouns, weak deficient pronouns and clitic pronouns) predict their availability to participate in FTIP competitions. The distribution, then, really relies on the FTIP itself. This approach is labeled by Safir as competitive narrowing.

The anaphoric forms treated in this chapter extend to include relational anaphors, such as SAME, OWN, BODYPART and OTHER. Safir furthermore provides some case studies of competitive narrowing in Greek, Malayalam, Japanese and Hungarian.

Chapter 7 is devoted to anaphora in the architecture of syntactic theory.
Having established a grammar that describes the behavior and interpretation of anaphors, Safir continues to look at the wider implications of the solutions he has posited. If his approach is the correct one, then what does that mean for the architecture of universal grammar?

Safir suggests dividing grammar into (a) a derivational module and (b) an interpretive module. The derivational module allows for convergent derivations to take place and should not be limited e.g. by minimalist effects, which have more to do with interpretation. It should, in other words, allow for “interpretation-blind derivation building”. The interpretive module compares convergent derivations and chooses the best one to represent a particular interpretation. The FTIP introduced by Safir is a prototypical example of what could be in this interpretive module. Safir mentions a few other principles that also compare convergent derivations. The first is called “Rule H” and ascertains that the shortest (c-command) distance to a possible antecedent should be taken (Fox 1998, 2000). The second is called “Weak Pronoun Competition” (WPC) and regulates the competition between more and less tonic pronouns (e.g. Samek-Lodovici 1995). The third one is the Independence Principle, which prohibits a constituent to be the antecedent for another constituent, if this second one c-commands it (Safir 2004).

Safir compares his two-level approach with existing one-level MP and OT approaches. He sees his two-level approach having practical advantages when compared to OT. The latter needs to process a lot of possibilities, when many constraints need to be taken into account. The former only needs to check a limited number of convergent derivations. Safir does briefly refer to bidirectional OT, which combines the form to meaning competition with the meaning to form one (Wilson 2001). What I haven’t found (but I may have missed it) is whether Safir’s own architecture also assumes bidirectionality. The FTIP algorithm assumes that someone who tries to interpret the meaning of an anaphor in a particular form (the listener) needs to put himself in the position of the speaker, in order to see what the effect would be if an alternative form had been chosen by the speaker.

**Conclusion**

Overall, I would say this book is a valuable asset for anyone who is interested in local anaphora. The reader should ideally be familiar with generative grammar—for the majority part of the book familiarity with the Minimalist Program is not necessary. Safir has managed to provide us with a theory that is capable of describing and resolving the intricacies of local anaphora. His general approach is convincing enough: there is competition between available forms, and there is a hierarchy of forms. Some elements I do not find convincing—notably the need to resort to covert grammar. The theoretical consequences that the success of the approach taken in this book takes should be taken seriously.

**References**


