IDENTITY AND INFORMATION:
SEMANTIC AND PRAGMATIC ASPECTS OF SPECIFICATIONAL SENTENCES

by

DAPHNA HELLER

A Dissertation submitted to the Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Linguistics

written under the direction of Professor Veneeta Dayal

and approved by

_____

VENEETA DAYAL

_____

JANE GRIMSHAW

_____

PAULINE JACOBSON

_____

ROGER SCHWARZSCHILD

New Brunswick, New Jersey

October, 2005

ABSTRACT OF THE DISSERTATION

Identity and Information: Semantic and Pragmatic Aspects of Specificational Sentences

By DAPHNA HELLER

Dissertation Director:
Veneeta Dayal

Specificational sentences are famous because they exhibit connectivity. Connectivity refers to the fact that the pre- and the post-copular phrase in a copular sentence behave as if they were in a c-command configuration. More Specifically, a number of syntactic and semantic phenomena that are usually found only under a c-command configuration are licensed in specificational sentences across the copula, i.e. where this configuration is absent.

This has lead researchers in the generative tradition to assume that a c-command configuration is available at an abstract level. This dissertation argues against such attempts, with special attention to the Question-Answer approach that has gained popularity in recent years. Instead, I argue for a direct compositional analysis of connectivity effects in which specificational sentences are analyzed as identity sentences as seen on the surface, and each connectivity effect is a by-product of certain syntactic and semantic combinatorics; the denotations of the elements are independently motivated for other environments. This analysis is applied to Hebrew, which has two subsets of specificational sentences, each exhibiting a different range of connectivity effects. It is

concluded that this pattern cannot be accounted for by any approach that derives connectivity using a single mechanism which characterizes specificational sentences.

Since there is no one-to-one correlation between connectivity and specificational sentences, it is concluded that connectivity is not a defining characteristic of specificational sentences. This leaves the question unanswered of what defines specificational sentences and how they differ from other copular sentences, especially from other identity sentences.

It is proposed that specification is defined by a certain pragmatic relation between the pre- and the post-copular expressions of an identity sentence. In particular, the post-copular expression is more 'discriminate' than the pre-copular expression with respect to the identity of the entity denoted. The notion of 'discriminability' is independently motivated by a new pattern from the choice of referring expressions. The discriminability perspective allows us to reduce the standard four-way typology of copular sentences into a three-way typology, and suggests that it may be possible to further reduce the typology to a two-way typology.

Acknowledgements

I would like to thank my advisor Veneeta Dayal and the other committee members Jane Grimshaw, Polly Jacobson and Roger Schwarzschild for their endless patience, their insightful questions and their thought-provoking comments.

For helping me at different stages of this work and in ever-so-many ways, I would like to thank Adrian Brasoveanu, Ivano Caponigro, Greg Carlson, Shai Cohen, Christine Gunlogson, Fred Landman, Heather Robinson, Susan Rothstein, Ivan Sag, Julie Sedivy, Rachel Sussman, Mike Tanenhaus, and, in particular, Yael Sharvit.

This work would not have been possible without the help of language consultants. I am indebted to Nurit Assayag, Rebecca Baier, Ivano Caponigro, Shai Cohen, Slavica Kochovska, Aniko Liptak, Orr Ravitz, Heather Robinson, Luca Storto, Rachel Sussman, Maryame Sy, Anna Szabolcsi, Harold Torrence and the native speakers of English at the UCLA linguistics department and at the University of Rochester.

Finally, I wish to thank my beloved family: Orr for always being there for me and Uriyah for making me smile.

Table of Contents

List of Tables

Chapter 1
Specificational Sentences: An Introduction

Specificational sentences are a subtype of copular sentences. Some examples of specificational sentences are given in (1-2). Intuitively, these sentences are called "specificational" because the post-copular phrase is a more specific description of whatever is mentioned in the pre-copular phrase. For instance, (1a) specifies that it is his tie that I don't like about John and (2a) specifies that Rome is the capital of Italy.

(1) a. What I don't like about John is his tie.
b. What I didn't find was any typos.
c. Where she had lunch was at the cafeteria.

(2) a. The capital of Italy is Rome.
b. The person John likes most is himself.
c. My next-door neighbor is Dan.

Specificational sentences have received much attention in the generative literature because they exhibit connectivity. Connectivity is the name for the effect that the two parts of a specificational sentence behave as if they were "connected" with respect to a number of syntactic and semantic phenomena. That is, specificational sentences license phenomena that would be expected if the material in the pre-copular phrase and the post-copular phrase were in a c-command configuration. For example, the Negative Polarity Item (NPI) *any* in the post-copular phrase in (1b) is licensed by the negation inside the pre-copular phrase and the post-copular reflexive in (2a) is bound by the name inside the pre-copular phrase. Both these phenomena are usually found when these elements occur in a c-command configuration. Although this configuration is not available here, the elements are still licensed – from this perspective, the two parts of the sentence are "connected". Connectivity effects are found for a number of syntactic and semantic

phenomena – we discuss the full range of the effects in section 1.

A less studied aspect of specificational sentences is the specificational relation itself. Higgins (1973) was the one who named these sentences 'specificational', discussing a range of their properties. While he presents no formal definition of the relation, he has provided an intuitive characterization by saying that the post-copular phrase "says what constitutes or makes up the object referred to by the subject noun phrase" (p. 150). A slightly more formal characterization comes from Akmajian (1970, p.19; cited in Higgins p. 153). Note that the term used by Akmajian is 'pseudo-clefts' – we come back to this term shortly.

> "The initial clause of the pseudo-cleft contains what is essentially a semantic variable, a semantic 'gap' which must be 'filled' or specified by the focus item [= the post-copular phrase DH]… The focus item must specify a value for the variable of the clause".

While these characterizations of the specification relation seem intuitively correct, the exact nature of specification is not well understood.

It has been suggested by Williams (1983) that specification is "inverse predication" (see also Partee 1986, Heggie 1988, Moro 1997 and Mikkelsen 2004b) – we come back to this approach in section 3. It should be pointed out here, however, that while this analysis is in line with the intuitive characterization of specification, it was never linked to connectivity.

A different analysis was proposed by Heycock & Kroch (1997, 1999) – they argue that specificational sentences are identity statements. Heycock & Kroch also present an analysis of connectivity: I briefly present their approach in section 1.1 below. If specificational sentences are simply identity sentences, then the two phrases around the copula have the same status. This symmetry does not capture the intuitive nature of

specification.

The only existing approach to specificational sentences that deals with both aspects is the Question-Answer approach which takes specificational sentences to be self-answering questions. This was originally proposed by Ross (1972) and recently developed by den Dikken, Meinunger & Wilder (2000) and by Schlenker (2003). In these analyses connectivity is derived in the obligatorily elided post-copular full IP answer. I discuss this approach in detail in chapters 2: I argue that there is no independent evidence to analyze specificational sentences as question-answer pairs. I will also examine it again in the light of the analysis presented in chapter 3.

This dissertation presents an analysis of both aspects of specificational sentences: the first part of the dissertation deals with connectivity and the second part considers the specificational relation in the light of what we will have concluded for connectivity. In discussing the two aspects we address questions about their possible relationship: can specificational sentences be defined as copular sentences that exhibit connectivity? Can specification be analyzed in such a way that will also account for connectivity? Or are these two phenomena independent?

The current chapter introduces the issues and gives the overall structure of the dissertation. But first, a note on the term 'pseudocleft' is in order. This term usually refers to a copular sentence introduced by a *wh*-word, such as the sentences in (1), repeated in part in (3a) for ease of reference. Since these sentences have a "corresponding" simple sentence, as illustrated in (3b), pseudoclefts have been considered to be a special construction, which is derived from the "corresponding" simple sentence by clefting. These are taken to be different from copular sentences that are not introduce by a *wh*-

word, as in (3c).

(3) a. What I didn't find was any typos.
b. I didn't find any typos.
c. The thing I didn't find was any typos.

Higgins (1973) has already pointed out that pseudoclefts are in essence no different from other copular sentences – they are copular sentences in which the pre-copular phrase is a free relative (we will see further evidence for this in chapter 2). This is expected if Free Relatives (FRs) are a type of nominals, as suggested by Jacobson (1995) (see also Caponigro 2003). As a result, we expect specificational pseudoclefts and specificational sentences that are not pseudoclefts to exhibit the same patterns – the next section shows that this is indeed what we find for connectivity. Therefore, in this dissertation the term 'pseudocleft' will be taken to mean 'copular sentence with a pre-copular FR' and will not be taken to be a separate construction. Note that under this interpretation of the term, pseudoclefts need not be specificational – they should be able to express any relation found in other copular sentences.

1. What Is Connectivity?

Since the monumental work of Higgins (1973), it is known that specificational pseudoclefts in English exhibit connectivity. Connectivity is the name for the effect that the two parts of a pseudocleft behave as if they were "connected", like the corresponding simple sentence. In particular, a cluster of syntactic and semantic phenomena that are usually assumed to hold only under a c-command configuration show up where this configuration is absent. One connectivity effect is the binding of a reflexive across the copula, illustrated in (4).

(4) a. John$_i$ is dangerous to himself$_i$
b. [What John$_i$ is _] is dangerous to himself$_i$

The reflexive in (4a) is assumed to be licensed as it is locally c-commanded by the antecedent *John*, respecting Principle A of the standard Binding Theory (Chomsky 1981). In the pseudoclefted counterpart (4b) the reflexive is not locally c-commanded by the antecedent, and yet the bound reading is grammatical. From the perspective of the standard Binding Theory, the two phrases around the copula are in some sense "connected": the reflexive behaves as if it were located in the gap inside the FR, although this gap is of course bound by *what*.

Note that not all pseudoclefts exhibit connectivity: these effects are found in specificational pseudoclefts, but not in predicational ones. Let us introduce predicational pseudoclefts by considering the ambiguous pseudocleft in (5).

(5) What John is is important.

On the predicational reading, *what John is* refers to John's job or position, and the sentence says that this job is important, e.g. if he is the president. The specificational reading talks about John himself, assigning the property *important* to him personally, which is very similar to simply saying *John is important*. Adding a reflexive to the post-copular phrase, as in (6), leaves only one reading – that John finds himself important, i.e. adding the reflexive renders the sentence unambiguously specificational.

(6) [What John$_i$ is _] is important to himself$_i$

(6) shows that the post-copular reflexive is licensed across the copula under the specificational reading but not under the predicational reading: this example illustrates that only specificational sentences exhibit connectivity. The two types of pseudoclefts also differ in other syntactic and semantic properties – we discuss these differences in detail in chapter 3. For now, we will assume that if a sentence exhibits connectivity it is specificational. While this is a rather common assumption, it is important that it is a

simplification because (i) some cases of connectivity are also found in predicational sentences, and (ii) as we will see in chapter 3, not all specificational sentences exhibit connectivity.

The rest of this section presents the full range of connectivity effects: what these phenomena have in common is that they are usually found under a c-command configuration, but in specificational sentences they occur where this configuration is absent. We will see that each of the connectivity effects is found both in a pseudocleft and in a non-pseudoclefted sentence – this provides preliminary evidence that pseudoclefts are not a special construction but rather a special case of copular sentences.

1.1. *Binding theory connectivity*

As we have seen above, this group of connectivity effects deals with the distribution of anaphoric elements. In the examples in (7), Principle A of the standard Binding Theory (Chomsky 1981) is violated: the anaphor *himself* is licensed in the post-copular phrase even though it is not c-commanded by the antecedent *John* which is embedded inside the pre-copular phrase. Despite the lack of a c-command configuration, the bound reading is grammatical.

(7) a. [What $John_i$ is _] is proud of $himself_i$.
b. [The person $John_i$ likes most _] is $himself_i$.

Principle B of the standard Binding Theory requires a pronoun to be locally free. In the examples in (8) the pronoun in the post-copular phrase is locally free, yet the coreferential reading is ungrammatical: the pronoun cannot take the nominal *John* as its antecedent as if it was c-commanded by it.

(8) a. *[What $John_i$ is _] is proud of him_i.
b. *[The person $John_i$ likes most _] is him_i.

The last connectivity effect in this group is Principle C connectivity. These are cases where a post-copular R-expression cannot corefer with a non-c-commanding pre-copular pronoun. In the examples in (9), *John* cannot be bound by *he*, even though it is locally free.

(9) a. *[What he_i is _] is a nuisance to $John_i$.
b. *[The people he_i saw _] were $John_i$ and some of Mary's friends. (Sharvit 1999)

The fact that I will be using the term 'Principle A/B/C' to talk about connectivity effects of anaphors should not be taken to imply that I adopt Binding Theory. We will see in chapter 3 that a non-configurational analysis of the distribution of anaphors is required in order to account for these connectivity effects.

1.2. *NPI connectivity*

Negative Polarity Items (NPIs) are also licensed across the copula. For example, the NPI *any* which is standardly assumed to be licensed only under the scope of negation occurs in the post-copular position in the sentences in (10) despite the fact that it is not in the scope of the pre-copular negation. (10a) illustrates NPI connectivity in a pseudocleft and (10b) illustrates the same connectivity effect in a non-pseudoclefted sentence. The ungrammatical sentences in (11) which differ minimally from (10) in that the pre-copular phrase is not negated indicate that it is the negation that licenses *any* in the post-copular position.

(10) a. [What John didn't buy _] was any books. (Sharvit 1999)
b. [The one thing he didn't do _] was buy any wine. (den Dikken et al. 2000)

(11) a. *[What John bought _] was any books.
b. *[The one thing he did _] was buy any wine.

1.3. *Opacity connectivity*

de dicto readings are usually available only in opaque contexts, i.e. under the scope of an intensional operator, where scope is defined in terms of c-command. Opacity connectivity, originally due to Halvorsen (1978), are cases where a post-copular nominal receives an opaque interpretation due to an intensional operator inside the pre-copular phrase.

In (12a), the post-copular phrase *a pink giraffe* has a *de dicto* reading, even though this phrase is not in the scope of the intensional predicate *look for* which is located inside the FR. That is, in addition to the *de re* reading which entails the existence of pink giraffes, this sentence has a reading which does not entail the existence of pink giraffes. Similarly, the intensional verb in the pre-copular phrase in (12b) licenses a *de dicto* reading for the post-copular phrase *a good job*, i.e. this sentence has a reading under which good jobs may not necessarily exist.

(12) a. [What John is looking for _] is a pink giraffe.
b. [The only thing that John is looking for _] is a good job. (Sharvit 1999)

1.4. *Bound variable connectivity*

The last connectivity effect we discuss is the binding of a pronoun by a quantifier across the copula. The usual configuration for a quantified expression to bind a pronoun is under a c-command configuration, but in (13), *no man* and *no student* bind *his* across the copula without being in a c-commanding position (both examples are from Sharvit 1999).

(13) a. [What [no student]$_i$ enjoys _] is his$_i$ finals.
b. [The women [no man]$_i$ listens to _] are his$_i$ wife and his$_i$ mother in law.

Bound Variable connectivity differs from the other connectivity effects discussed here in that it is also found in predicational sentences – such effects are discussed in Williams

(1994), Sharvit (1997) and Heycock & Kroch (1999). These effects are illustrated in (14) where the post-copular pronoun is bound by the quantifier inside the pre-copular FR.

(14) a. What [every student]$_i$ got was a nuisance to him$_i$.
b. What [every man]$_i$ is is a nuisance to him$_i$.

In (14a), *every student* binds the post-copular *him* – the sentence has a reading where the thing(s) every student got has the property of being a nuisance to that student, e.g. if John who hates sports got running shoes and Bill who hates camping got a tent. Similarly, (14b) has a reading where the (contextually) unique property that every man has is a property that annoys that man, e.g. if John who is hates impatience is impatient and Bill who hates arrogance is arrogant.

For further discussion of Bound Variable connectivity, see Sharvit (1999). What is relevant for our purposes here is the fact that Bound Variable connectivity is found in predicational sentences – this means that unlike the previous effects we saw that are only found in specificational sentences, Bound Variable connectivity cannot be used as a diagnostic for whether a copular sentence is specificational. Note that this answers one question we have presented in the beginning: specificational sentences cannot be defined as copular sentences that exhibit connectivity, since connectivity is found in sentences that are not specificational.

1.5. *Summary*

We have discussed four types of connectivity effects in English, but additional effects are found in other languages. In the next chapter we will see connectivity effects in Hungarian, Wolof, Macedonian and Hebrew. In chapter 3, we will see in Hebrew two connectivity effects that are not found in English: subject-predicate agreement and Case connectivity. While connectivity effects may be very different from each other, they do

share one property: all these phenomena are usually found under a c-command configuration, but in specificational sentences they are found where this configuration is absent – this is what stands behind the name "connectivity".

The reason connectivity has received much attention in the literature is that its analysis has implications that go beyond copular sentences. If the standard analyses of the different grammatical phenomena do not apply in specificational sentences where they show up as connectivity effects, these analyses cannot be maintained. There are two possible ways to go. The first is to preserve these standard analyses which make use of c-command and to analyze these effects in specificational sentences by postulating the relevant c-command configuration at some abstract level. The other option is to take connectivity effects as evidence against using c-command in the analyses of these different grammatical phenomena and to develop new analyses which do not make use of this structural notion. The next section reviews the two approaches in more detail.

2. Analyses of Connectivity

2.1. *The c-command approach*

The first analysis of connectivity was proposed by Peters & Bach (1968). Peters & Bach posits a level of representation at which the post-copular phrase is surrounded by a copy of the FR. In a similar approach, developed in Hornstein (1984), the c-command relation is achieved by having the material from the post-copular phrase in the position of the gap inside the FR. The two analyses are illustrated in (15): (15a) is a specificational pseudocleft (repeated from 6) that exhibits Principle A connectivity; in (15b) the material inside the FR was copied to the post-copular position; in (15c) the post-copular material was copied to the gap inside the FR.

(15) a. [What John$_i$ is _] is important to himself$_i$
b. [What John$_i$ is _] is John$_i$ is important to himself$_i$
c. [What John$_i$ is important to himself$_i$] is important to himself$_i$

These analyses face a number of problems – see Higgins (1973, chapter 2) for the earliest discussion. But the main objection to deriving connectivity in this way is that the level of representation in which the c-command relation is posited makes little semantic sense. If we copy the post-copular materials into the position of the gap, the pre-copular phrase will be neither a well-formed nominal nor a well-formed sentence. If, on the other hand, we copy material from the pre-copular phrase around the post-copular phrase, the two phrases around the copula will no longer be of the same type – recall that the requirement that the pre- and the post-copular phrases be of the same type was already mentioned by Akmajian. Since, as pointed out by Jacobson (1994), the abstract level is the one at which binding relations are licensed, the representations should be semantically coherent. Moreover, we will see in chapter 3 that a complex connectivity pattern like the one in Hebrew cannot be accounted for by such mechanisms.

A more recent mechanism for deriving connectivity was proposed in Heycock & Kroch (1999). They posit a post-LF level of representation where the sentence undergoes "iota-reduction": the logical representation of a specificational pseudocleft in (16a) is manipulated to yield the corresponding simple sentence, as in (16b).

(16) a. [What John$_i$ likes _] is himself$_i$
b. (ιx: John likes x) = himself $\rightarrow$ John likes himself

This approach faces two problems. The conceptual problem, pointed out in Schlenker (2003), is that Heycock & Kroch do not present independent evidence for positing the additional level of representation: it would be preferable to account for connectivity without positing a level of representation that has no independent evidence. More

seriously, this approach is only suited for sentences that have a "corresponding" simple sentence, i.e. pseudoclefts. This approach cannot account for the same connectivity effects in non-pseudoclefted specificational sentences, because if the logical representation of (17a) is as in (17b), it cannot be manipulated using this transformation as there is no "corresponding" simple sentence.

(17) a. [The person $John_i$ likes _] is $himself_i$.
b. (ιx: x is a person & John likes x) = himself → ???

If iota-reduction can only account for a subset of connectivity effects, i.e. for connectivity effects in pseudoclefts, we will need a different account for connectivity effects in non-pseudoclefted specificational sentences. This is not a desirable result as connectivity does not seem to be different in these two types of sentences, and, in addition, if such an account is developed, it may account for all connectivity effects rendering iota-reduction superfluous.

The last analysis that aims to account for connectivity by positing the desired c-command at an abstract level is the Question-Answer analysis, originally due to Ross (1972) and recently developed by den Dikken, Meinunger & Wilder (2000) and by Schlenker (2003). This analysis takes specificational sentences to be a pair of a question and its answer – c-command is posited in the post-copular full answer[1]. This analysis is more attractive than the ones mentioned above because it derives connectivity effects using mechanisms that already exist in the grammar: the reason for reconstruction is not

[1] Romero (2003) is also identified with this line of analysis, as her work draws on similarities of the pre-copular phrase in a specificational sentence and (concealed) questions. Her analysis, however, provides a choice between an individual concept and a clausal analysis for the question-answer equation and hence is merely compatible with deriving connectivity by using a c-command configuration in the post-copular position. We discuss Romero's analysis in more detail in the next chapter.

particular to specificational sentences. Nicely, this is the same as Question-Answer pairs. This is illustrated in (18): (18a) shows how Principle A connectivity is derived in a question-answer pair and (18b) applies this analysis to a specificational pseudocleft showing Principle A connectivity.

(18) a. What is John? John is proud of himself.
b. [$_{Q}$ What John is] is [$_{ANS}$ John is proud of himself].

In addition to providing an account for connectivity, the question-answer approach is also attractive because, as noticed out by Higgins (p. 202), specificational sentences are similar in meaning to question-answer pairs: the specification given in the post-copular phrase for what constitutes the pre-copular phrase is similar to answering a question. Since this is the most promising version among the c-command restoring analyses, it will be the focus of chapter 2.

Chapter 2, which is based on joint work with Ivano Caponigro, argues against analyzing the pre-copular phrase in a specificational sentence as a question, both syntactically and semantically. We first show that the pre-copular phrase in a specificational sentence is not an interrogative, i.e. is not a question syntactically, contra the proposal of den Dikken, Meinunger & Wilder (2000). We show that although FRs and embedded *wh*-interrogatives have the same form in English, there are distributional differences between the two constructions. In particular, the range of *wh*-words that occur in FRs is a subset of those found in embedded *wh*-interrogatives. Crucially, this is the same subset of *wh*-words that occur in the pre-copular *wh*-clause of specificational pseudoclefts. To demonstrate that this is not an accidental correlation, we turn to languages that distinguish FRs and embedded *wh*-interrogatives morphologically. We show that in these languages only FRs can occur in a specificational pseudocleft. The

evidence we present comes from four genetically unrelated languages – Macedonian (a Slavic language), Hungarian (a Finno-Ugric language), Wolof (a Niger-Congo West Atlantic language) and Hebrew (a Semitic language).

The second part of chapter 2 argues against analyzing the pre-copular phrase in a specificational sentence as a semantic question, i.e. as a syntactic nominal that is interpreted like a question. It is known since the work of Baker (1968) that some nominals in English can receive such an interpretation in certain environments – these are called "concealed questions". Schlenker (2003) proposes that the pre-copular position of a specificational sentence hosts concealed questions (Romero 2003 also draws on similarities to concealed questions, but see again fn. 1). We present three arguments against this analysis. First, we show that not all languages that exhibit connectivity also allow for concealed questions, and thus, concealed questions cannot be responsible for the existence of connectivity. Macedonian provides an example for such a language in that it does not allow for concealed question nominals in canonical environments and yet it does exhibits connectivity in specificational sentences. Second, we show that not all the nominals that occur in the pre-copular position of specificational sentences can occur in the canonical concealed question environment – this would be unexpected if these nominals are freely interpreted as concealed questions. Finally, we compare the expected interpretation of FRs in the canonical concealed questions environment with their interpretation in the pre-copular position of specificational sentences: if the two environments trigger the concealed question interpretation, we expect the two to be the same – but they are not.

The reason for focusing on the status of the pre-copular phrase is that it serves as the

motivation for analyzing the sentence as a question-answer pair. If there is no independent evidence that the pre-copular phrase is a question in some sense, the question-answer analysis is merely a stipulation and loses it attractiveness. One could of course also examine the post-copular phrase to see whether it is an answer. But since an answer is a declarative sentence, being an answer depends on the discourse status of the sentence and not merely on its syntactic or semantic properties. Thus, it would be harder to provide conclusive evidence regarding the answerhood status of the post-copular phrase.

Furthermore, answers can take different forms – full or short, so in examining the status of the post-copular phrase one has to consider both options. This is especially relevant since the goal of the question-answer approach is to account for connectivity in specificational sentences. In particular, the analyses of den Dikken et al. (2000) and Schlenker (2003) crucially depend on the status of the post-copular phrase as an (obligatorily) elided answer in order to account for connectivity in the standard c-command configuration which is available in the full answer. Using a full answer to account for connectivity faces two problems. First, there are cases where the full IP exhibits the opposite pattern of what we find in specificational sentences – these are the so-called "anti-connectivity" effects (Sharvit 1999; see also Schlenker 2003). Second, there are cases where there is no obvious full IP which will have the desired c-command. Both these problems will be discussed in chapter 2.

Note, interestingly, that showing that connectivity cannot be accounted for by positing a post-copular full IP does not rule out the possibility that specificational sentences are question-answer pairs. In particular, specificational sentences can be

analyzed as self-answering questions in which the post-copular phrase is a short answer – this is illustrated in (19) which is parallel to (18).

(19) a. What is John? proud of himself.
b. [$_{Q}$ What John is] is [$_{ANS}$ proud of himself].

Of course, pursuing this version of the question-answer analysis requires a different account of connectivity that does not rely on c-command. While this is a logical possibility, it should be noted that such an analysis has not been presented in the literature.

Importantly, the arguments presented in chapter 2 argue against any version of the question-answer approach, including Question-short Answer, because they are aimed at the status of the pre-copular phrase. If the pre-copular phrase is not a question, the motivation for the question-answer approach is lost, independent of the exact status of the answer. So while chapter 2 discusses the specific analyses presented in den Dikken et al. (2000) and in Schlenker (2003) in which connectivity is accounted for in the post-copular full IP answer, the arguments are relevant for any version of the question-answer theory.

2.2. *The direct compositional approach*

The alternative approach to connectivity takes the heterogeneity of connectivity effects as a sign that connectivity is not actually a single grammatical phenomenon. Rather, this approach takes connectivity effects to show that the analyses we currently have for the different grammatical phenomena that show up as connectivity effects are incorrect and should be revised to not rely on c-command. While the approaches we saw in the previous section wish to maintain the standard analyses for the different grammatical phenomena and get these to apply in specificational sentences by positing abstract structure which contains the necessary c-command configuration, the goal of the direct

compositional approach is to license all grammatical phenomena without using abstract structure. Therefore, this approach requires developing new analyses for the phenomena that would apply in all the environments, i.e. without using c-command which is not available in all cases. The direct compositional approach is reminiscent of Higgins' (1973) "null hypothesis" (p. 13): "the surface structure form of a specificational pseudo-cleft sentence is essentially identical to its deep structure form".

This approach, which was nicknamed "the semantic theory" because it does not rely on the syntactic notion of c-command in analyzing the different grammatical phenomena, was developed by Jacobson (1994) and Sharvit (1999) (also see Cecchetto 2000,2001). They take specificational sentences to express identity and propose new analyses for the different grammatical phenomena such that they apply in all environments, including specificational sentences. Connectivity comes out as a by-product of the combinatorics of identity semantics. For example, Principle A connectivity in (20) is analyzed by Jacobson (1994) as identity between the (unique) property that John has and a reflexive predicate – this amount to saying that John is proud of himself.

(20) a. What John is is proud of himself.
b. $\iota P_{<e,t>}[P_{<e,t>}(j_e)] = \lambda x_e.\text{proud-of}'(x_e,x_e)$

Chapter 3 takes these analyses as its starting point and develops an account of the connectivity pattern in Hebrew using direct compositional analyses. The connectivity pattern in Hebrew is especially interesting because not all specificational sentences exhibit connectivity. In particular, Hebrew has two types of specificational sentences: those where the (impersonal) pronominal copula agrees with the post-copular phrase, and those where it has a fixed "neutral" form (which is identical in form to the masculine singular agreeing copula). Interestingly, the form of the copula affects the availability of

connectivity effects. The examples in (21) demonstrate that opacity connectivity shows up on the (feminine) post-copular NP only when the copula is neutral (nouns in Hebrew are marked for gender).

(21) [ma še-dan mexapes _] ze / zot jirafa vruda
what that-Dan seeks is(n) / is(f) giraffe(f) pink(f)
'What Dan seeks is a pink giraffe.'

Although the post-copular nominal is not c-commanded by the intensional verb, the neutral copula pseudocleft is ambiguous between a *de dicto* reading (where the existence of pink giraffes is not entailed) and a *de re* reading (where the existence of pink giraffes is entailed). This ambiguity is not found when the copula agrees with the post-copular feminine NP: the sentence only has the *de re* reading. I will argue that this difference stems from the type of the arguments that the two copulas equate: the agreeing copula is restricted to individuals, while the neutral copula can equate any type of arguments. Opacity connectivity is analyzed following Sharvit (1999): the ambiguous pseudocleft (with the neutral copula) will be an equation of properties and the de re only pseudocleft will be an equation of individuals.

What makes the connectivity pattern in Hebrew especially complicated is the fact that not all connectivity effects are absent from pseudoclefts with the agreeing copula. One connectivity effect that is found with both types of copulas is Bound Variable connectivity, illustrated in (22).

(22) mi še-[kol gever]$_i$ ohev ze /zot ima šel-o$_i$
who that-every man loves is(n)/is(f) mother of-his
'The person every man loves is his mother.'

Both the neutral copula version and the agreeing copula version allow for the bound reading, i.e. the sentence is true if every man loves his own mother. I follow Jacobson (1994) and Sharvit (1999) and use functions to derive the bound reading in these cases:

both types of pseudoclefts are analyzed as an equation of <e,e> functions. This type is argued to be possible with the agreeing copula as it is a formal extension of type e denotations.

Chapter 3 deals with all connectivity effects in Hebrew, some of which have not been discussed before, such as Case marking and agreement connectivity. The connectivity pattern in Hebrew shows that there is no one-to-one correlation between specificational sentences and connectivity. This means that connectivity should not be seen as a defining characteristic of specificational sentences. The "mixed-bag" pattern in Hebrew cannot be analyzed by any theory that derives all connectivity effects by a single grammatical operation and thus indicates that connectivity is not a single phenomenon. In conclusion, the Hebrew connectivity pattern provides conclusive evidence that each connectivity effect needs a separate analysis.

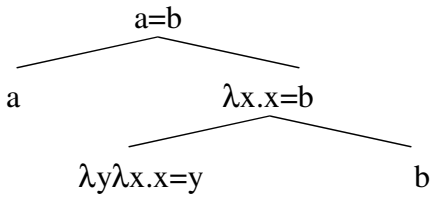
Taking together chapters 2 and 3, we conclude that connectivity is best explained as a by-product of the combinatorics of equation rather than a result of any abstract structure or hidden operation. But if specificational sentences are what we see on the surface – copular sentences that express identity, we do not have an explanation for the special nature of the specification relation. The second part of the dissertation addresses this issue using a new notion of discriminability.

3. THE SPECIFICATIONAL RELATION

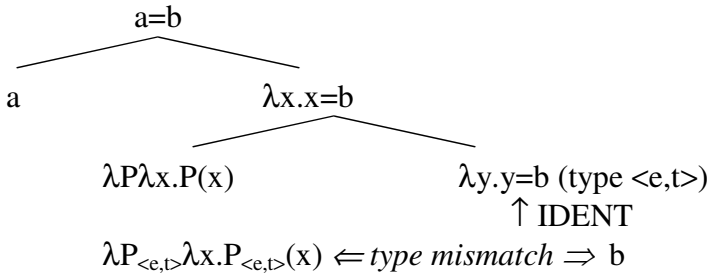
The direct compositional analysis of connectivity that will be presented in chapter 3 depends on specificational sentences expressing identity at the sentence level. Specificational sentences will express identity at the sentence level if they are identity statements. However, if the grammar contains Partee's (1987) IDENT type-shifting rule,

identity at the sentence level can also be derived as a special case of predication or inverse predication. In particular, the denotation of an individual can be shifted to denote the singleton predicate which will be predicated over the individual. (23a) illustrates the composition of identity directly; (23b) illustrates the composition of identity as a special case of predication; (23c) illustrates the composition of identity as a special case of inverse predication. Note that the copula in the cases of predication and inverse predication has the same meaning, but it takes its arguments in the opposite order – this meaning for the copula has been proposed in Partee (1986).

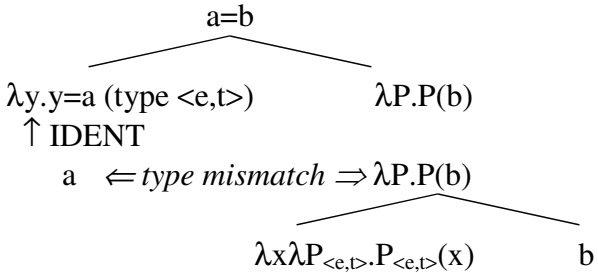
(23) a. Identity



b. Identity as a special case of predication



c. Identity as a special case of inverse predication



Therefore, the analysis of connectivity advocated in chapter 3 of this dissertation does not provide a conclusive answer regarding the nature of specificational sentences. However,

since we will see in chapter 3 that predicational and specificational sentences behave differently with respect to a number of structural tests, we can conclude that specificational sentences do not express identity as a special case of predication. We are left with two options: direct identity and identity as a special case of predication.

Analyzing specificational sentences as inverse predication would provide an answer for what specification is. Independent of connectivity, Williams (1983) proposed that in a specificational sentence the post-copular phrase is the subject of predication and the pre-copular phrase is predicated over it. This proposal was first formalized by Partee (1986) and was later given other formalizations by Heggie (1988) and Moro (1997) and recently by Mikkelsen (2004b). If the post-copular semantic subject is an instance of the property denoted by the pre-copular predicate, it can viewed as "specifying" it. However, Heycock & Kroch (1997, 1999) have argued that English specificational sentences are not inverse predication and in previous work (Heller 1999) I have argued the same for Hebrew.

Instead, Heycock & Kroch argue that specificational sentences are direct identity sentences. However, this analysis does not provide any explanation as to the nature of the specificational relation. On the contrary: the relation of identity implies some symmetry between the arguments, but such specificational sentences are actually asymmetric. While I maintain that the semantic relation between the arguments in specificational sentences is indeed identity, I propose that the asymmetry is found along a different dimension. In particular, I introduce a pragmatic notion of 'discriminability' and propose that the two expressions around the copula differ in their discriminability. That is, they differ in the level of familiarity with the entity they reflect. In other words, they differ in how well they identify the entity they denote, which is translated into how well they can

discriminate it from other entities.

Note, in conclusion, that the question-answer approach also provides an analysis of the specificational relation by drawing on similarities with questions and their answers. However, since we will see in chapter 2 that the pre-copular phrase in a specificational sentence is in no sense a question, we will not consider this analysis in the light of the direct compositional analysis of connectivity.

3.1. *Introducing discriminability*

Chapter 4 presents a new pattern from the choice of referring expressions and accounts for this pattern as a scalar phenomenon along the dimension of 'discriminability'. We will see that if the context is such that there is information in the common ground that allows constructing more than one referring expression for the same entity, the expression that must be used is the more 'discriminate' one. In other words, if the interlocutors are familiar with certain aspects of the entity, the speaker has to mention these aspects in the referring expressions even if they are not necessary for establishing successful reference.

As an illusttation, consider (24) in the context of a potluck party held at my house.

(24) Q Who brought the lasagna?
A1 Giacomo.
A2 The next-door neighbor.

You and I are eating some of the delicious lasagna, and you inquire about the identity of the person who brought the lasagna. In a context where we both know the neighbor's name, i.e. the common ground contains information that allows constructing both expressions for the same individual, the two answers can both pick out the same relevant individual for both speakers. Nonetheless, the proper name constitutes an appropriate

answer, while using the definite description is odd (changing the name to a person familiar to the reader might help here). This is unexpected if the only goal in using a referring expression is to pick it out in order to say something about it.

Of course, if one of the interlocutors is not aware of the equi-extensionality of the expressions, the expression that must be used is the shared one. That is, this pattern of preference is only found if the relevant information is available in the common ground. In addition, this pattern is only relevant for expressions used referentially, in the sense of Donnellan (1966). If the speaker wishes to pick out the individual by virtue of its being her next-door neighbor, the description will be appropriate.

I discuss preferences among four types of expressions: proper names, definite descriptions with contentful nouns, definite descriptions with "bleached" nouns, and free relatives. All are illustrated in (25) which deals with reference to location.

(25) Q Where did Dan go on his vacation?

A1 To London.
A2 To the capital of England.
A3 To the place he went last year.
A4 Where he went last year.

I discuss this pattern with reference to individuals, locations and objects in three languages: English, Italian and Hebrew. I also discuss how other expressions relate to this pattern: indefinites, pronouns and *-ever* FRs. While we will see that these are not directly relevant for the pattern, they shed more light on this phenomenon.

The obvious question is whether this pattern reflects a preference of certain linguistic forms over others or whether the preference is for certain information, and the preference for forms are a by-product of the information that is usually encoded in certain forms. We address this question and conclude that it is the information that is relevant.

The preference pattern is analyzed as a scalar phenomenon. In particular, Grice's Maxim of Quantity requires the cooperative speaker to give all the information available to her. Here this requirement concerns information about her assumed familiarity with the entity. This is communicated to the addressee indirectly via the choice of the referring expression. If the less discriminate expression available is used, it gives rise to a scalar implicature; it implicates that a more discriminate expression was not available in the common ground. Note, importantly, that the scalar implicature does not arise if the addressee can find a reason for violating Grice's Maxim in the form of an extra message, e.g. avoiding the neighbor's name in (24) in order to signal that I dislike him: this would be a flouting implicature.

This phenomenon shows that referring expressions also have a secondary role. In particular, they are used indirectly by the interlocutors to make sure that they are "on the same page". More formally, to maintain a non-defective context in the sense of Stalnaker (1978), i.e. to keep their common grounds the same.

3.2. *Discriminability in copular sentences*

Chapter 5 uses the notion of discriminability argued for in chapter 4 to examine the specificational relation. I argue that while a specificational sentence asserts that the two phrases around the copula denote one and the same entity, what makes these sentences specificational is that the post-copular phrase is a more discriminate expression than the pre-copular expression. Consider as an illustration the unambiguously specificational pseudocleft in (26).

(26) What I don't like about John is his tie.

Why does this sentence lack a predicational reading? There should be a reading where *his tie* is predicated of the entity which is picked out by the pre-copular free relative, i.e. the thing I don't like about John. Higgins' explanation for the unavailability of the predicational reading is that this free relative is inherently non-referential and hence cannot occur in the pre-copular position of a predicational sentence. This, however, cannot be the explanation because, as (27) shows, the same FR is possible in this position when the predicate is adjectival: this sentence is appropriate if I am giving you hints about the nature of the thing I don't like about John.

(27) What I don't like about John is dotted.

Similarly, we can demonstrate that the post-copular phrase from (26) can occur in the post-copular position of a predicational sentence. I therefore conclude that the wellformedness of a copular sentence depends on the relation between the expressions, and propose that the relevant dimension for the relation is discriminability.

For specificational sentences, I propose that they are identity sentences in which the post-copular expression is more discriminate than the pre-copular expression. Identity is also available as a special case of predication (see again 23b). In these cases, the opposite discriminability relation is observed: the post-copular expression is more discriminate than the pre-copular expression. So the reason (26) is unambiguously specificational is due to the relative discriminability of the expressions around the copula. Note, crucially, that decreasing discriminability does not characterize predicational sentences in the same way that rising discriminability characterizes specificational sentences. This is because discriminability is only relevant for copular sentences in which the pre- and the post-copular expressions denote the same entity. In predicational sentences, this is not always the case.

Taking this proposal as the starting point, we examine how other specificational sentences fit with this pattern. We will see that discriminability is not an absolute property of the expressions. First, the discussion of proper names in copular sentences will show that discriminability has to be relativized to the specific discourse context. Second, discussing copular sentences with two headed descriptions will show that discriminability has to be relativized both to the entity denoted: the same expression can be more or less discriminate depending on how it relates to the nature of the entity. While investigating the notion of discriminability in copular sentences shows that it is a more fine grained notion than originally assumed for referring expressions, I will not at this point propose a definition of discriminability that would make the necessary distinctions. In addition, we will see that the distinctions observed for the choice of referring expressions cannot always be carried over to copular sentences. Nonetheless, the discriminability analysis sheds new light on the analysis of copular sentences.

Since the discriminability analysis of specificational sentences is reminiscent of the inverse predication analysis in that it takes predicational and specificational sentences to be in some sense the opposite, we will compare the two. We will see that the predictions of the inverse analysis are incorrect with respect to the range of the possible specificational sentences and with respect to reversed specificational sentences. The identity + discriminability analysis, on the other hand, can deal with these constructions.

Examining copular sentences from the perspective of discriminability also sheds light on the typology of copular sentences. First, we will see that this perspective on copular sentences allows seeing the similarities between predicational and identificational sentences. I will argue that Higgins' four-way typology of copular sentences can be

reduced into a three-way typology, distinguishing predicational, specificational and identity sentences. I will also propose that a better understanding of discriminability may allow further reducing the copular typology, distinguishing predicational and identity sentences and deriving other types using discriminability.

Chapter 2

The Question-Answer Approach*

As mentioned in the previous chapter, one of the primary approaches to specificational sentences is the question-answer approach. This approach, originally due to Ross (1972), takes specificational sentences to be a pair of a question and its answer. Recently, it has been given two implementations. Den Dikken, Meinunger & Wilder (2000) analyze the *wh*-clause in a specificational pseudocleft as an embedded interrogative (the term *interrogative* will be use in this chapter to refer to a syntactic sentence type – it is more restricted than *question*). Schlenker (2003) proposes that the pre-copular phrase is syntactically a nominal, but that this nominal is interpreted as a question (this analysis has been called 'Question-in-Disguise', a term due to Sharvit 1999). In both analyses, connectivity is accounted for by restoring the desired c-command configuration in the (obligatorily elided) post-copular answer.

This approach to specificational sentences is attractive because the semantics of these sentences is intuitively similar to that of questions and answers. For example, one characterization of specificational sentences proposed by Higgins (1973) is that "the subject in some way delimits a domain and the specificational predicate identifies a particular member of that domain" (p.213). This is intuitively similar to the role an answer plays of filling in the variable in the question. Moreover, this approach does not introduce new notions into the grammar – it uses notions that are used to analyze question-answer pairs.

---

* This chapter is based on a joint paper with Ivano Caponigro that was presented at the "Workshop on Direct Compositionality", Brown University, June 19th 2003 – see Caponigro & Heller (to appear).

This chapter presents crosslinguistic data arguing that the pre-copular phrase in specificational sentences is not a question, neither syntactically nor semantically. The arguments in this chapter have been developed in collaboration with Ivano Caponigro. The reason for focusing on the status of the pre-copular phrase is that it stands at the core of the question-answer approach: specificational pseudoclefts show similar effects to question-answer pairs independent of whether the answer is full or short. That is, connectivity effects are found in question-answer pairs whether the answer is full or short. Hence, while both specific analyses we discuss in this chapter belong to the 'Question plus Deletion' family (this term is due to Romero 2003), there could be a version of the question-answer approach that accounts for connectivity effects in question-answer pairs where the answer is short, i.e. not a full IP. Therefore, in order to argue against all possible versions of the question-answer analysis we need to show that the pre-copular phrase is in no sense a question.

The chapter is organized as follows. Section 1 introduces the question-answer approach to specificational sentences. Section 2 argues against den Dikken et al.'s (2000) version of the question-answer approach by showing that, crosslinguistically, the *wh*-clause in the pre-copular position of a specificational pseudocleft is not an embedded *wh*-interrogative, but rather a Free Relative (FR). Section 3 argues against Schlenker's (2003) version of the question-answer approach which takes the pre-copular phrase in a specificational sentence, whether a FR or a headed nominal, to be interpreted as a question. We present crosslinguistic data showing that the pre-copular phrase in a specificational sentence is not a Concealed Question (CQ), and discuss the implications of our findings to the study of CQs. Section 4 deals with the status of the post-copular

phrase – we present anti-connectivity effects and discuss how they affect the analysis of connectivity within the question-answer approach.

1. Connectivity in the Question-Answer Approach

The resemblance of the *wh*-clause in specificational pseudoclefts in English to an embedded *wh*-interrogative has led Ross (1972, 1997) to propose that specificational pseudoclefts are an equation between a question and an obligatorily elided answer, which is assumed to be derived from the full answer by a process of phonological deletion. To support the existence of question-answer equations, Ross presents examples in which the post-copular answer is not elided.

(1) a. What I did then was [call the grocer]. (Ross 1972)
b. What I did then was [I called the grocer].

(2) a. What John did was [buy some wine]. (den Dikken et al. 2000)
b. What John did was [he bought some wine].

The logic is that in order to account for the existence of (1b) and (2b), one has to assume that the grammar allows for question-answer pairs in copular sentences. Having a pair of a question and an elided answer, as in (1a) and (2a), comes "for free" due to the independent existence of ellipsis in answers. That is, these examples show that analyzing specificational sentences as question-answer pairs does not involve postulating new mechanisms in the grammar. This is an advantage of the question-answer approach over any account of connectivity that involves a special copying or deletion mechanism – see again chapter 1.

One main advantage of this analysis is the account of connectivity effects it allows for: the effects are licensed by c-command in the post-copular full answer. This is illustrated in (3-8) for the connectivity effects we saw in the previous chapter. The (a)

examples are question-answer pairs and the (b) examples are the corresponding specificational pseudoclefts.

(3) *Principle A*
a. What is John? ~~John$_i$ is~~ proud of himself.
b. [$_{\mathbf{Q}}$ What John is] is [$_{\mathbf{ANS}}$ ~~John is~~ proud of himself].

(4) *Principle B*
a. What is John$_i$? *~~John$_i$ is~~ proud of him$_i$.
b. *[$_{\mathbf{Q}}$ What John$_i$ is] is [$_{\mathbf{ANS}}$ ~~John$_i$ is~~ proud of him$_i$].

(5) *Principle C*
a. What is he$_i$? *~~He is~~ proud of John$_i$.
b. *[$_{\mathbf{Q}}$ What he$_i$ is] is [$_{\mathbf{ANS}}$ ~~he is~~ proud of John$_i$].

(6) *NPI*
a. What didn't John buy? ~~John didn't buy~~ any books.
b. [$_{\mathbf{Q}}$ What John didn't buy] was [$_{\mathbf{ANS}}$ ~~John didn't buy~~ any books].

(7) *Opacity*
a. What is John looking for? ~~John is looking~~ ~~for~~ a pink giraffe.
b. [$_{\mathbf{Q}}$ What is John looking for] is [$_{\mathbf{ANS}}$ ~~John is looking for~~ a pink giraffe].

(8) *Bound Variable*
a. What does [no student]$_i$ enjoy? ~~[No student]$_i$ enjoys~~ his$_i$ finals.
b. [$_{\mathbf{Q}}$ What [no student]$_i$ enjoys] is [$_{\mathbf{ANS}}$ ~~[No student]$_i$ enjoys~~ his$_i$ finals].

Note that the fact that analyzing specificational pseudoclefts (and, more generally, specificational sentences) as question-answer pairs allows for a straightforward account of connectivity does not by itself constitute evidence in favor of this analysis – it is only a motivation to look for independent evidence.

This chapter examines the status of the pre-copular phrase in specificational sentences and shows that there is no evidence that it is in any sense a question. The next section shows that the *wh*-clause in a specificational pseudocleft is syntactically a FR and not a *wh*-interrogative, contra what has originally been proposed by Ross (1972) and adopted by den Dikken et al. (2000). Then, in section 3, we show that the pre-copular phrase in a

specificational sentence is also not interpreted like a question, contra Schlenker's (2003) version of the question-answer approach.

2. Question-Answer I: The Embedded Interrogative Approach

Den Dikken, Meinunger & Wilder (2000) adopt Ross's original idea and analyze specificational pseudoclefts as "self-answering questions". In their analysis, the pre-copular *wh*-clause is an embedded *wh*-interrogative and the post-copular phrase is an (obligatorily) elided full IP which answers the question in the pre-copular *wh*-interrogative. Their primary motivation comes from Ross' examples in which the post-copular IP is not elided, as we saw in the previous section. Note, however, that the existence of question-full answer pairs as we saw in (1a) and (2a) only shows that such copular sentences are allowed by the grammar. It does not show that specificational sentences are such question-answer pairs.

Den Dikken et al. distinguish this type of pseudoclefts (which they call "Type A") from reversed pseudoclefts ("Type B") which they analyze as predicational sentences in which the predicate is a FR. This distinction is motivated by the irreversibility of NPI connectivity. In particular, they notice that NPIs are licensed in the post-copular phrase by the negation inside the pre-copular *wh*-clause, as in (9a) and (10a), but reversing the order of the elements around the copula renders the sentences ungrammatical, as in the (b) examples, i.e. the NPI in the pre-copular phrase is not licensed by the negation inside the post-copular *wh*-clause.

(9) a. What John didn't buy was any books. **Type A**
a'. [What John didn't buy] was [$_{IP}$ ~~he didn't buy~~ any books].
b. *[$_{DP}$ Any book] is/was [what John didn't buy]. **Type B**

(10) a. What wasn't available was a doctor who knew anything about babies.
a'. [What wasn't available] was [$_{IP}$ t~~here wasn't available~~ a doctor who knew anything about babies].
b. *[$_{DP}$ A doctor who knew anything about babies] was [what wasn't available].

Taking as their starting point the standard assumption that NPIs are licensed by a c-commanding negation, den Dikken et al. assume that the licensing of an NPI in the (a) examples indicates that negation is present in a c-commanding position, and hence conclude that the post-copular phrase is an elided full IP. Den Dikken et al. claim that the ungrammaticality of the (b) examples suggests that the pre-copular phrase inType B pseudoclefts is not an elided IP but rather an XP – a DP in the examples here.

This section is not intended to review den Dikken et al.'s arguments. Instead, it presents crosslinguistic data arguing that the *wh*-clause in a specificational pseudocleft is not an embedded interrogative, but rather a FR. Our logic is that if the pre-copular phrase is not a question, the post-copular phrase cannot be the answer, so we lose the motivation for reconstructing a full IP in the post-copular position.

2.1. *The range of wh-words in free relatives and embedded interrogatives*

Although the *wh*-clause in a specificational sentence may look identical to an embedded *wh*-interrogative, a closer examination reveals distributional differences between the two constructions. In particular, the range of *wh*-words that occur in FRs is a subset of those found in embedded *wh*-interrogatives. Crucially, this is the same subset of *wh*-words that occur in the pre-copular *wh*-clause of specificational pseudoclefts. Note that we are comparing the *wh*-clause in a specificational sentence to embedded *wh*-interrogatives because matrix *wh*-interrogatives in English have a different form: they exhibit Subject-Aux inversion and *do*-support. However, if we compared the *wh*-clause in a specificational sentence to matrix *wh*-interrogatives, the same argument would hold.

The examples in (11) present the range of *wh*-words in embedded interrogatives. In (12), the same *wh*-clauses are used in the complement position of non-interrogative predicates, i.e. in FRs: the FRs introduced by *what* (12a) and *where* (12b) are grammatical, while the FR introduced by *who* (12c) is marginal and these introduced by complex *wh*-expressions such as *which*+NP or *how much* (12d,e) are ungrammatical (for more on the range of *wh*-words in FRs, see Caponigro 2003). The specificational pseudoclefts in (13) exhibit the exact same restrictions on *wh*-words as FRs.

(11) *Embedded Interrogatives*
a. I wonder [*where* she has lunch].
b. I wonder [*what* John is reading].
c. I wonder [*who* gave you the flowers].
d. I wonder [*which bo*ok John is reading].
e. I wonder [*how much* Sue weighs].

(12) *Free Relatives*
a. I have lunch [*where* she has lunch].
b. I read [*what* John is reading].
c. ??I met [*who* gave you the flowers].
d. *I read [*which book* John is reading].
e. *I weigh [*how much* Sue weighs].

(13) *Specificational Pseudoclefts*
a. [*Where* she has lunch] is at the cafeteria.
b. [*What* John is reading] is “Ulysses”.
c. ??[*Who* gave you the flowers] was your advisor.
d. *[*Which book* John is reading] is “Ulysses”.
e. *[*How much* Sue weighs] is 130 pounds.

These data show that although in English the *wh*-clause in specificational pseudoclefts is morphologically identical to an embedded *wh*-interrogative, their distribution is different: *wh*-clauses in specificational pseudoclefts patterns with FRs and not with interrogatives.

It should be pointed out that *-ever* FRs are irrelevant for this discussion, since, as pointed out by Jacobson (1995), they are banned from specificational pseudoclefts (but see Dayal 1997 who shows that they can occur in specificational sentences with a post-

copular FR. We will come back to this in chapter 5). This is exemplified in (14). Hence, the fact that they differ from standard FRs in that they allow for the same range of *wh*-words as embedded interrogatives, as exemplified in (15), does not provide additional evidence.

(14) *Specificational Pseudoclefts*
a. *[Wherever she has lunch] is at the cafeteria.
b. *[Whatever John reads] is "Ulysses".
c. *[Whoever gave you the flowers] was your advisor.
d. *[Whichever book John reads] is "Ulysses".
e. *[However much Sue weighs] is 130 pounds.

(15) *-ever Free Relatives*
a. I will have lunch [wherever she has lunch].
b. I will read [whatever John reads].
c. I want to meet [whoever gave you the flowers].
d. I will read [whichever book John is reading].
e. I want to weigh [however much Sue weighs].

Going back to standard FRs, it should be noted that den Dikken et al. do mention (albeit in a footnote) the difference between the range of *wh*-words in embedded interrogatives and in specificational pseudoclefts. They, however, attribute this facts to a restriction on the kinds of interrogatives that can appear in specificational pseudoclefts. Unfortunately, they do not offer any insight as to what this restriction may be, so at this point it is merely a stipulation. However, since their proposal is logically possible, we turn to languages that distinguish FRs and *wh*-interrogatives on the surface for more evidence.

2.2. *Morphological differences between wh-interrogatives and specificational pseudoclefts*

This section examines the *wh*-clause in specificational pseudoclefts in four languages – Macedonian, Hungarian, Wolof and Hebrew – and provides morphological evidence that the *wh*-clause in the pre-copular position of a specificational pseudocleft is a FR and not

an embedded interrogative. Following Caponigro (2003), we will use the complement position of two types of predicates to distinguish the two – interrogative predicates for embedded interrogatives and argument predicates for FRs (for more on distinguishing FRs and interrogatives, see Caponigro 2003, p. 12).

In Macedonian, FRs differ from embedded interrogatives in that they are introduced by *ona* 'that'. When the *wh*-clause occurs in the complement of *kazhi* 'tell', as in (16a), *ona* cannot occur, and when the same *wh*-clause occurs in the complement of *sakam* 'love', as in (16b), *ona* must occur. Crucially, *ona* is also obligatory in the specificational pseudocleft in (16c): this pseudocleft is made sure to be specificational as it exhibits Principle A connectivity.

(16) MACEDONIAN

a. *Embedded Interrogative*
Kazhi mi (***ona**) shto navistina Petar saka
tell me that what really Petar love
'Tell me what Petar really loves.'

b. *Free Relative*
(Jas) sakam *(**ona)** shto Petar saka
I love that what Petar loves
'I love what Petar loves.'

c. *Specificational Pseudocleft*
*(**Ona)** shto Petar saka e samiot sebe si
that what Petar loves is alone himself
'What Petar loves is himself.'

In Hungarian, the words that introduce FRs are characterized by a prefix *a-* that makes them distinguishable from the *wh*-words that introduce interrogatives. In (17a) the *wh*-clause occurs as the complement of *mondd* 'tell', i.e. it is an interrogative, and in (17b) the *wh*-clause occurs as the complement of an individual taking verb, i.e. it is a FR. While in the former environment only *mit* 'what' is possible, the opposite pattern is

observed in the latter environment where only *amit* 'what' can occur. Crucially, only *amit* can occur in the specificational pseudocleft in (17c) – this sentence exhibits opacity connectivity; it is ambiguous between a *de dicto* and a *de re* reading.

(17) HUNGARIAN

a. *Embedded Interrogative*
Mondd meg [***amit/mit** fo"zött]
tell me $what_{FR}$/$what_{INT}$ cooked
'Tell me what he cooked.'

b. *Free Relative*
Megettem [**amit/*mit** fo"zött]
I-ate $what_{FR}$/$what_{INT}$ cooked
'I ate what he cooked.

c. *Specificational Pseudocleft*
[**Amit/*mit** keres _] az Chomsky legújabb könyve
$What_{FR}$/$What_{INT}$ is-looking-for that C.'s latest book
'What he is looking for is Chomsky's latest book.'

In Wolof, a Niger-Congo West Atlantic language spoken mainly in Senegal and Gambia, the (contracted) *wh*-words result from combining the many classifiers of the language with the suffix *-u*, while the words that introduce FRs are formed by adding the suffix *-i* to the same classifiers. Again, we compare the clause that occurs as the complement of an interrogative taking verb like *yëg* 'found out' in (18a) with that of an individual taking verb like *bañ* 'hate' in (18b). The former environment only allows for a clause introduced by *l-u* while the latter requires a clause introduced by *l-i*. Crucially, the specificational pseudocleft in (18c), which exhibits Principle A connectivity, allows only for *l-i*, i.e. this is a FR and not an interrogative (*cl-FR* is a classifier + free relative morpheme; *cl-INT* is a classifier + interrogative morpheme).

(18) WOLOF

a. *Embedded Interrogative*
yëg -na [***l-i /l-u** móódu gën-ë bëgg]
find out-neutral cl-FR/cl-INT Moodu surpass-inf like
'She found out what Moodu likes most.'

b. *Free Relative*
bañ-na [**l-i /*l-u** móódu gën-ë bëgg]
hate-neutral cl-FR/cl-INT Moodu surpass-inf like
'She hates what Moodu likes most.'

c. *Specificational Pseudocleft*
[**l-i /*l-u** móódu gën-ë bëgg _] bopp-am la
cl-FR/cl-INT Moodu surpass-inf like head-3sgposs be
'What Moodu likes most is himself.'

In Hebrew, FRs are distinguished morphologically from *wh*-interrogatives in that they require the occurrence of the complementizer *še*. In the complement position of the verb *berer* 'inquired' in (19a), the *wh*-clause cannot contain the complementizer. In the complement position of the verb *kara* 'read' in (19b), the complementizer *še* must occur. The specificational pseudocleft in (19c) patterns with (19b) in that it requires occurrence of the complementizer – this sentence is ensured to be a specificational pseudocleft as it exhibits both Principle A connectivity and Case connectivity (for more on connectivity in Hebrew, see chapter 3).

(19) HEBREW

a. *Embedded Interrogative*
dan berer [ma (***še**)-karati] (Sharvit 1999)
Dan inquired what that$_{COMP}$-(I)-read
'Dan inquired what I read.'

b. *Free Relative*
dan kara [ma *(**še**)-karati] (Sharvit 1999)
Dan read what that$_{COMP}$-(I)-read
'Dan read what I read.'

c. *Specificational Pseudocleft*
[ma *(**še**)-dan ohev _] ze et acmo
what that$_{COMP}$-Dan loves is Acc himself
'What Dan loves is himself.'

Den Dikken et al. (2000) also mention languages that show a similar pattern to what is presented here for Hebrew, Wolof, Hungarian and Macedonian (their footnote 23). In particular, they cite Bulgarian (following Izvorski 1997) and Greek (following Alexiadou & Giannakidou 1998) as languages that distinguish interrogatives and FRs overtly and employ only the latter in specificational pseudoclefts. Den Dikken et al. propose analyzing these cases as their "Type B" pseudoclefts, i.e. as simple copular sentences that do not involve questions and answers. The same analysis can be applied to the languages discussed here. But this would leave us with six languages (mostly genetically unrelated) in which den Dikken et al.'s analysis does not apply. That is, even if there are languages in which the *wh*-clause in a specificational sentence is an embedded interrogative, as proposed by den Dikken et al., this is not true of specificational pseudoclefts crosslinguistically and is therefore not a general account of connectivity.

Going back to English, for which den Dikken et al. proposed their analysis, we face an additional problem. Recall from the previous chapter (chapter 1, section 1) that connectivity is also found in specificational sentences in which the pre-copular phrase is a headed nominal and not a *wh*-clause. In these cases, it is not clear that the pre-copular phrase is an interrogative in any syntactic sense. Den Dikken et al. propose that these nominals are in fact elided embedded interrogatives. For example, they propose that the non-pseudoclefted specificational sentence in (20a) is derived from (20b).

(20) a. The one thing he didn't do was buy any wine.
b. [$_{CP}$ ~~What~~ [the one thing he didn't do] ~~was~~ *t*] was [~~he didn't~~ buy any wine].

As noted by the authors, pursuing this analysis requires an explanation "for why ellipsis of this sort … is restricted to the 'topic' questions of specificational pseudoclefts" (p. 83). In other words, the authors acknowledge that the suggested ellipsis is highly specialized: the only questions that are elided in this way are copular questions and this ellipsis is only found in the pre-copular position of specificational sentences. While this may turn out to be a necessary kind of ellipsis, den Dikken et al. do not present any independent evidence that this is indeed the case – at this point, their suggestion is nothing more than a stipulation.

In sum, den Dikken et al.'s (2000) version of the question-answer approach where the pre-copular phrase in a specificational sentence is analyzed as an embedded interrogative is only applicable to a very limited number of cases of connectivity effects. First, it does not apply to languages where the *wh*-clause is clearly not an embedded interrogative – we have mentioned six such languages: Macedonian, Hungarian, Wolof, Hebrew, Bulgarian and Greek. Second, this analysis only applies to pseudoclefts, but, as we saw in the previous chapter, the full range of connectivity effects is also found in copular sentences with a pre-copular headed nominal. The way den Dikken et al. propose for dealing with these sentences is stipulative. Another necessary stipulation of this analysis is the restriction on the *wh*-words that are possible in pseudoclefts. The number of stipulations required in this analysis suggests that it is not in the right direction.

The next section turns to consider a version of the question-answer approach that avoids the problems raised here by analyzing the pre-copular phrase in a specificational sentence as a syntactic nominal and assuming that this nominal is interpreted as a question.

3. Question-Answer II: The Concealed Question Approach

Schlenker (2003) acknowledges that the *wh*-clause in a specificational pseudocleft is syntactically not an interrogative, but rather a FR. To maintain the question-answer analysis, he proposes that these FRs, as well as all other nominals that occur in the pre-copular position of specificational sentences, are interpreted as questions. Interpreting the pre-copular phrase as a question motivates positing a post-copular answer. In Schlenker's analysis, this answer is an (obligatorily elided) full IP in which the desired c-command configuration is available.

The crucial part is how to motivate a question interpretation for a nominal. Schlenker proposes that these are Concealed Questions (CQ), in the sense of Baker (1968). The canonical environment for CQ nominals is the complement position of (certain) interrogative taking predicates. This is illustrated in (21-24) where the (a) examples are embedded interrogatives and the (b) examples are the corresponding CQs (Baker p.81).

(21) a. Jane figured out [$_{CP}$ what the plane's arrival time is].
b. Jane figured out [$_{DP}$ the plane's arrival time].

(22) a. John refused to tell the police [$_{CP}$ who the fellows who has been involved were].
b. John refused to tell the police [$_{DP}$ the fellows who has been involved].

(23) a. Susan found out [$_{CP}$ what the place where the meeting was to be held is].
b. Susan found out [$_{DP}$ the place where the meeting was to be held].

(24) a. Fred tried to guess [$_{CP}$ what the amount of the stolen money was].
b. Fred tried to guess [$_{DP}$ the amount of the stolen money].

While the complements in both versions are interpreted as questions, they take different syntactic forms. Grimshaw (1979) argues that this phenomenon shows that the semantic selection (s-selection) of predicates is independent from their syntactic, or categorial, selection (c-selection) (see Dor 1992 for a different view). We can thus say that the predicates in (21-24) s-select for questions and c-select for either DPs or CPs which

yields two possible combinations. Schlenker's proposal then is that the pre-copular position of a specificational sentence is another case of a CQ environment. In other words, he proposes that while the copula c-selects for DPs, it may also s-select for questions (note that the selection here is for the pre-copular position as opposed to the complement position in the case of the predicates above).

This proposal is not radical: the idea that the copula is cross-categorial has been in the literature since Montague (1973) (see also Partee 1986). The goal of this section is not to determine whether there exist question-answer pairs in copular sentences, but rather to assess the claim that specificational sentences are such question-answer pairs. In arguing against this analysis, we again focus on the status of the pre-copular phrase. This is because the CQ interpretation is the motivation reconstructing a post-copular full IP answer.

The section is organized as follows. Section 3.1 presents Schlenker's arguments for interpreting the pre-copular phrase in a specificational sentence as a CQ. Section 3.1. briefly discusses Romero's (2003) version of the question-answer approach, showing that Romero's arguments do not constitute evidence in favor of Schlenker's account of connectivity. Section 3.3 presents a global argument against associating the availability of connectivity effects with CQ interpretations. Section 3.4 points out distributional differences between the pre-copular position of specificational sentences and the canonical CQ environment and section 3.5 discusses interpretative differences between the two. Finally, section 3.6 evaluates the implications of the data presented here for the study of CQ nominals.

3.1. *Schlenker's (2003) arguments*

Schlenker presents a number of arguments in favor of his question-answer analysis, two of which concern the status of the pre-copular phrase. Since our discussion focuses on the status of the phrase in this position, we only present these two arguments here.

Before presenting the arguments themselves, it should be noted that the data Schlenker presents come from sentences that exhibit connectivity in French – these sentences are an instance of left dislocation rather than simple copular sentences. Schlenker does not address the question of whether these are specificational sentences and instead proposes that his analysis applies to a class he calls 'connectivity sentences', which he defines by the existence of connectivity effects. This classification is problematic because crosslinguistically the class of sentence that exhibit connectivity does not seem to form a coherent natural class. One exception in the class of 'connectivity sentences' are predicational sentences that exhibit Bound Variable connectivity (see again chapter 1 section 1). Crucially, these sentences exhibit only one connectivity effect, but if we adopt Schlenker's classification they are expected to exhibit either all or no connectivity effects. In addition, in the next chapter we will see that in Hebrew there are specificational sentences that exhibit only a subset of connectivity effects. Since I believe Schlenker intends his analysis to apply crosslinguistically, I will assume here that it should apply to specificational sentences rather than to 'connectivity sentences'.

Schlenker's first argument concerns the form of 'connectivity sentences' in French. Schlenker claims that the fact that these sentences require a pronominal element, as illustrated by the contrast in (25), shows that the pre-copular phrase is not an individual,

as this is the pattern in identity sentences of non-individuals, as shown by the contrast in (26).

(25) a. ??Ce qu'il aimait était lui-même FRENCH
it that-he liked was himself
b. Ce qu'il aimait, c'était lui-même
it that-he liked it-was himself
'What he liked was himself'.

(26) a. ??Être vieux est être oublié
to-be old is to-be forgotten
b. Être vieux, c'est ennuyeux
to-be old that-is boring
'To be old is boring.'

The problem with this argument is that we are missing a crucial example of individual identity that would show that the same pronominal is not required in identity sentences of individuals (Schlenker provides the predicational version of 26, but it does not complete the pattern). But even if it is indeed the case that identity sentences of individuals do not require the pronominal element, then all this argument shows is that the pronominalization facts are <u>compatible</u> with analyzing the dislocated phrase as denoting a question. This is not a necessary result and hence is only a weak argument in favor of the question-answer analysis.

Schlenker presents a stronger argument that comes from question anaphora. Consider first the connectivity sentence in (27a) which exhibits Binding connectivity – this is shown by the comparison to the simple sentence in (27b).

(27) FRENCH
a. Ce qu'il aimait c'était lui-même/??lui/*l'imbécile/*Jean
it that-he liked it-was himself /him /the idiot /Jean
'What he liked was himself / him / the idiot / Jean.'

b. Il aimait lui-même/??lui/*l'imbécile/*Jean
he liked himself /him /the idiot /Jean
'He liked himself / him / the idiot / Jean.'

In (28) the same nominal occurs in the complement position of the interrogative-taking verb *demandé* 'ask' – the pronominal in the coordinated sentence exhibits the same connectivity effects.

(28) FRENCH

je me suis longtemps demandé [ce qu'il aimait] et j'ai finalement appris
I me am long asked it that-he liked and I-have finally learned
que c'était lui-même/??lui/*l'imbécile/*Jean
that it-was himself /him /the idiot /Jean
'For a long time I have been wondering what he liked, and have finally learned that it was himself/him/the idiot/Jean.'

This shows that connectivity effects arise when the pre-copular phrase has a question denotation. However, the environment considered here is not that of specificational sentences. Thus, while it does allow us to conclude that a question denotation may license connectivity effects, it does not show that this is the way connectivity is licensed in specificational sentences. That is, as acknowledged by Schlenker himself, this is only an indirect argument in favor of the question-answer analysis.

The rest of this section evaluates the claim that the pre-copular phrase in a specificational sentence is a CQ by examining data from five languages. But before presenting these arguments, we briefly present Romero's (2003) version of the CQ approach to specificational sentences showing that it does not support Schlenker's analysis.

3.2. *Romero's (2003) version of the concealed question approach*

Like Schlenker, Romero (2003) argues that the pre-copular phrase in a specificational sentence has the same interpretation as a CQ nominal (see also Romero, in press). She bases her analysis on two semantic similarities between nominals occurring in the two positions. First, she notices that certain ambiguities discussed by Heim (1979) for CQs

are also found in the pre-copular position of specificational sentences. Consider the ambiguity of the CQ nominal in (29).

(29) John knows the price that Fred knows.
Reading A. John knows the same prices that Fred knows.
Reading B. John knows which price Fred knows.

Both readings require a context where there are several relevant questions about prices, e.g. 'How much does milk cost?', 'How much does bread cost?' 'How much does butter cost?', and in which Fred knows the answer to one of these, e.g. the first one. Under reading A, John knows the answer to the same question, i.e. he also knows the price of milk. Under reading B, in contrast, John doesn't know the answer to this question, but rather knows which question it is that Fred knows the answer to, i.e. he knows that Fred knows the price of milk – that is, he knows an answer to a meta-question about the price question. Romero points out that the same ambiguity is found with nominal occurring in the pre-copular position of specificational sentences, as illustrated in (30).

(30) a. The price that Fred thought was $1.29 was actually $1.79.
Reading A. The question whose answer Fred thought was $1.29 was actually $1.79.

b. The price that Fred thought was $1.29 was the price of milk.
Reading B. The question whose answer Fred thought was $1.29 was 'how much is the price of milk?'.

These readings are found in similar contexts, i.e. there are several questions about prices and Fred knows the answer to one of them. The sentence in (30a) which illustrates reading A says that Fred mistakenly thought that the price of milk was $1.29, but in fact it was $1.79. The sentence in (30b) illustrates reading B which concerns the meta-question, namely, which price it is that Fred thought was $1.29.

The second similarity Romero notices between nominals occurring in the two environments concerns the particle *also*. In particular, she notes that in these

environments the nominals receive a non-exhaustive reading, in contrast with the exhaustive reading found, for example, in the pre-copular position of predicational sentence. To illustrate, consider first the underlined subject in (31). The sentence asserts that everything Carlos was wearing, with the exception of the nice hat, didn't suit him – this is an exhaustive reading.

(31) *Pre-copular phrase in a predicational sentence*
A: I heard Carlos was wearing a very nice hat yesterday that everybody admired.
B: What he was also wearing didn't suit him at all. And, on top of that, it was really expensive.
A: I hadn't heard anything about that…

This contrasts with the non-exhaustive reading of the same phrase in the specificational sentence in (32). If the same phrase in (32) had an exhaustive reading, it would mean that Carlos was only wearing a hat and a pair of tight pants, but the sentence does not have this reading. The CQ nominal in (33) is also non-exhaustive: here again the sentence does not mean that Carlos was half naked.

(32) *Pre-copular phrase in a specificational sentence*
A: I heard Carlos was wearing a very nice hat yesterday that everybody admired.
B: What he was also wearing was a pair of tight orange pants that didn't suit him at all. And, on top of that, they were really expensive.
A: I hadn't heard anything about that…

(33) *CQ nominal*
A: I heard Carlos was wearing a red hat yesterday that everybody admired.
B: Do you want me to tell you the garment he was also wearing? A pair of tight orange pants that didn't suit him at all. And, on top of that, they were really expensive.
A: I hadn't heard anything about that…

Romero proposes a unified analysis for nominals occurring in the pre-copular position of specificational sentences and in canonical CQ environments. These analyses are formally complex and I will not go into the details here. What is important for us here is that Romero proposes two different accounts and does not choose between them.

Importantly, in one account the denotation of the nominals is of the type of individual concepts <s,e>, whereas in the other it is of the propositional type <s,t>. This dual analysis shows that one can account for the similarities between CQ nominals and the pre-copular phrase in specificational sentences without having to assume that these denote propositions. If the pre-copular phrase is not propositional in some sense, there is no motivation for a post-copular full answer, in which case we lose the analysis of connectivity proposed by Schlenker (because connectivity effects are accounted for in the full answer).

Therefore, Romero's work cannot be used to argue for or against a question-answer analysis à la Schlenker: it leaves the question open whether specificational sentences are an equation of propositions or individual concepts, whereas for Schlenker it is crucial that the equation would be of propositions[1].

3.3. *Argument I: concealed questions are not responsible for connectivity*

Our first argument against the question-answer analysis of specificational sentences does not deal with the status of the pre-copular phrase in a specificational sentence directly but rather examines the more general connection between CQ nominals and connectivity. Since Schlenker's version of the question-answer approach ties the availability of connectivity effects to the availability of CQ interpretations, we would expect any language that exhibits connectivity in specificational sentences to have some CQ

[1] For Schlenker, specificational sentences are crucially clausal equations. To this end, his semantics of the construction uses Groenendijk & Stokhof's (1990, 1997) semantics for questions in which the extension of a question is a proposition. However, Schlenker also proposes an alternative implementation using the Karttunen (1977) / Hamblin (1973) line in which questions are sets of propositions.

nominals. Contrary to this prediction, Macedonian exhibits connectivity but does not allow for CQ nominals.

First, as we saw in section 2.2, Macedonian has specificational sentences that exhibit connectivity. (34) presents two kinds of connectivity effects.

(34) MACEDONIAN

a. *Principle A connectivity*
[Ona shto Petar saka _] e samiot sebe si.
that what Petar loves is alone himself
'What Petar loves is himself.'

b. *Opacity connectivity*
[Ona shto Petar bara _] e najnovata kniga od Chomsky
that what Petar look-for is latest-the book by Chomksy
'What Petar is looking for is Chomsky's latest book.'

However, no nominals in Macedonian can be interpreted as CQs. (35) shows that even nominals that are easily interpreted as CQs in other languages (see section 3.4 below) do not receive a CQ interpretation in Macedonian. When these nominals occur in the complement position of the Macedonian equivalent of the predicate *kazhi* 'tell', the resulting sentences are totally unacceptable[2].

(35) MACEDONIAN

a. *Kazhi mi go {saatot / chasot / vremeto}.
tell me it hour-the / hour-the / time-the
('Tell me the time.')

b. *Kazhi mi ja {tezhinata / tvojata tezhina}.
tell me it weight-the / your-the weight
('Tell me your weight.')

[2] This predicate was chosen because the correlates of *tell* crosslinguistically seem to be the most permissive in allowing CQ nominals. We have checked some other predicates in Macedonian that can take interrogative complements – *prashuva* 'wonder', *otkrie* 'discover' and *doznava* 'found out' – but these do not allow for CQ nominals either.

c. *Kazhi mi ja {goleminata / tvojata golemina} na chevlite.
tell me it size-the / your-the size-the of shoes-the
('Tell me your shoe size.')

The fact that connectivity is found in a language that does not allow for CQ nominals in the canonical CQ environment suggests that the two phenomena are unrelated. It is still logically possible that Macedonian has CQ interpretation which is specific to the pre-copular position of specificational sentences. We examine this possibility, albeit for English, in section 3.5.

3.4. *Argument II: distributional differences*

If the pre-copular nominal in specificational sentences, whether a FR or a headed nominal, receives a CQ interpretation, we expect correlations with the canonical CQ environment. This section presents data demonstrating that FRs and some headed nominals that occur in the pre-copular position in specificational sentences cannot occur in the canonical CQ environment, i.e. the expected correlations are not found. This suggests that the pre-copular position in a specificational sentence does not in fact receive a CQ interpretation.

3.4.1. Free relatives

If FRs are freely interpreted as questions in the pre-copular position of specificational sentences (i.e. in pseudoclefts), we expect that they would also occur in the canonical CQ environment, i.e. as complements of interrogative predicates. As pointed out by Sharvit (1999), this prediction can only be tested in languages that distinguish FRs and *wh*-interrogatives morphologically and also allow for CQ nominals. Sharvit tests this prediction for Hebrew, which allows nominals to occur in the complement position of interrogative predicates and to receive a CQ interpretation (36a). Not surprisingly,

embedded *wh*-interrogatives can freely occur in this position (36b), but, crucially, FRs cannot. Recall from section 2.2. (example 19) that FRs in Hebrew differ from *wh*-interrogatives in the presence of the complementizer *še*.

(36) HEBREW

a. *Concealed Question*
dan berer et [$_{DP}$ ha-sha'a]
Dan inquired Acc the-hour
'Dan inquired about the time.'

b. *Embedded Interrogative*
dan berer [ma karati]
Dan inquired what (I)-read
'Dan inquired what I read.'

c. *Free Relative*
*dan berer [ma **še**-karati]
Dan inquired what that$_{COMP}$-(I)-read
'Dan inquired what I read.'

In section 2.2 we saw three other languages that distinguish embedded interrogatives and FRs overtly: Macedonian, Wolof and Hungarian. In the previous section (section 3.3) we saw that Macedonian does not allow for any CQ nominals, so examining FRs is irrelevant here. In the rest of this section we apply Sharvit's argument to Wolof and Hungarian. These languages show the same pattern as Hebrew.

In Wolof, the verb *birëlé* 'find out' can take as its complement a CQ nominal in (37a) and an embedded interrogative in (37b) but not a FR in (37c): the two clausal arguments differ in the word that introduces them: l-u for interrogatives and l-i for FRs.

(37) WOLOF

a. *Concealed Question*
móódu birëlé-na [$_{DP}$ waxtu-wu ñëw-u avioŋ bi].
Moodu find.out-NEUTRAL time-u arrive-u airplane the
'Moodu found out the airplane's arrival time.'

b. *Embedded Interrogative*
birëlé-na [**l-u** móódu gën-ë bëgg].
find out- NEUTRAL cl-INT Moodu surpass-INF like
'She found out what Moodu likes most.'

c. *Free Relative*
* birëlé-na [**l-i** móódu gën-ë bëgg].
find out-NEUTRAL cl-FR Moodu surpass-INF like
'She found out what Moodu likes most.'

Turning to Hungarian, we again find CQs in the complement of an interrogative taking verb, as in (38a). The same environment of course allows for *wh*-interrogatives, as in (38b), but not for FRs, as in (38c): the two are distinguished morphologically in the form of the *wh*-word.

(38) HUNGARIAN

a. *Concealed Question*
Mondd meg [$_{DP}$ az eredményt].
tell me the score
'Tell me the score.'

b. *Embedded Interrogative*
Mondd meg [**mit** fo"zött].
tell me what$_{INT}$ cooked
'Tell me what he cooked.'

c. *Free Relative*
*Mondd meg [**amit** fo"zött].
tell me what$_{FR}$ cooked
'Tell me what he cooked.'

The data presented here show that FRs cannot occur in the canonical position of CQs. If FRs freely received a CQ interpretation as proposed by Schlenker, this would be an unexpected result. It is of course still possible that FRs receive a CQ interpretation only in the pre-copular position of a specificational pseudocleft, i.e. via a context sensitive mechanism. We argue against this option in section 3.5. But, first, we turn to distributional facts concerning headed nominals.

3.4.2. Headed nominals

The previous section examined the distribution of FRs in the canonical CQ environment and in the pre-copular position in specificational sentences. This section does the same for headed nominals – we find that some nominals that occur in the pre-copular position of specificational sentences and thus, according to Schlenker, should have a CQ interpretation cannot occur in the canonical CQ environment. (39a) is a specificational sentence with the lexical nominal *the president of the United States* in the pre-copular position, which can also occur in the complement position of an interrogative predicate. By contrast, an apparently similar individual-denoting nominal like *the boy who ran over my pet snake* can occur in the specificational sentence in (39b), but not in the canonical CQ environment in (39b'). (40) and (41) show the same contrast for different lexical items.

(39) a. [The president of the United States] is G.W. Bush.
a'. Tell me [the president of the United States].
b. [The boy who ran over my pet snake] was John.
b'. */??Tell me [the boy who ran over my pet snake].

(40) a. [The capital of France] is Paris.
a'. Tell me [the capital of France].
b. [The city I live in] is Paris.
b'. ??Tell me [the city you live in].

(41) a. [The candy Jill wants to buy] is jelly beans.
a'. Tell me [the candy Jill wants to buy].
b. [The money that was stolen] was Swiss Franks.
b'. *Tell me [the money that was stolen].

These data show that the expected correlation between the pre-copular position of specificational sentences and canonical CQ environments is not found. Like with FRs, it is possible that the nominals in the (b) examples are not freely interpreted as CQs, but

they can receive such interpretation in the pre-copular position of specificational sentences. Unfortunately, this claim seems to be untestable.

3.5. *Argument III: interpretation differences*

We in saw in section 3.4.1 that FRs do not receive a CQ interpretation in the canonical CQ environment and hence concluded that they do not receive such an interpretation via a general context insensitive mechanism. However, it is still possible that the pre-copular position of a specificational sentence is special in that it allows for a CQ interpretation of FRs and other nominals that occur in this position. This section examines this claim directly by looking at the interpretation of FRs in this position. We will see that the interpretation FRs receive in this position is different from the expected interpretation of FRs as a CQs.

If FRs do receive a CQ interpretation in specificational sentences, the question arises as to what this interpretation would be. In order to answer this question, we examine the interpretation of the relevant string in the canonical CQ environment. Consider, for example, the interpretation of the *wh*-clause in (42a) – we expect it to be similar to the interpretation of the nominal in (42b).

(42) a. Tell me [what the capital of France is _].
b. Tell me the capital of France.

But what does the *wh*–clause mean in (42a)? This sentence is asking to identify Paris. That is, it would be fine to reply to (42a) by saying *Paris*, but it would be totally infelicitous to reply with *beautiful*.

Our next step looks at the interpretation of this string in a specificational pseudocleft. Interestingly, in this context we find the opposite pattern. In particular, if the same *wh*-clause occurs in the pre-copular position of a specificational pseudocleft, the post-

copular phrase must be a property like *beautiful* and not an individual like *Paris*:

(43) a.*[What the capital of France is _] is Paris.
b. [What the capital of France is _] is beautiful.

If we compare the two environments we see that the FRs in (43) get a different interpretation from what is expected from (42): while the CQ asks for an individual, the specificational sentences requires a post-copular property. That is, even when we examine the interpretation of a FR in the pre-copular position of a specificational we do not find the CQ interpretation. This pattern shows that the pre-copular position of a specificational sentence does not receive a CQ interpretation even via a context sensitive mechanism.

3.6. *A note on concealed question nominals*

The data presented in sections 3.4-3.5 was used to argue that the pre-copular position in specificational sentences is not interpreted as a CQ. However, this data is also relevant to the study of CQ nominals. In particular, we saw that not all nominals can occur in the canonical CQ environment. The banned nominals were FRs and nominals headed by certain lexical nouns. The kinds of nominals that can occur in the canonical CQ environment has not been discussed in the CQ literature – this literature is mostly concerned with characterizing the predicates that allow for CQ nominals (Grimshaw 1979, Heim 1979, Dor 1992). An analysis of CQs is beyond the scope of this dissertation, but, since this has not been discussed before, it is worth noting which nouns are possible in the canonical CQ environment. Therefore, we also propose a preliminary generalization. Our hope is that this will contribute to future research on CQ nominals.

The examples we saw in section 3.4.2. contrasted the CQ nominals *president*, *capital* and *candy* with the non-CQ nominals *boy*, *city* and *money*. (44-47) present examples of

nouns that can occur in the canonical CQ environment (as marked, some of the examples are cited from previous work).

(44) a. John found out the **murderer** of Smith. (Heim 1979)
b. Tell me the **president** of the United States.
c. Tell me the **chair** of your department.
d. Tell me the **winner** of last year's Pulitzer Prize.
e. Tell me the **writer** who won the last Pulitzer Prize.

(45) a. John discovered the **location** of the meeting. (Dor 1992)
b. Tell me the **capital** of France.

(46) a. John knows Bill's **telephone number**. (Heim 1979)
b. Harold guessed the **time** of the meeting. (Dor 1992)
c. Tell me your shoe **size**.
d. Tell me your **height**.
e. I couldn't figure out her **age**.
f. Guess the **temperature** of the water.
g. Tell me the **amount** of money that was stolen.
h. Please tell me the **grade** you got in that class.

(47) a. Harold knew the **kind** of candy that Jill liked. (Dor 1992)
b. Harold learned the **outcome** of the trial. (Dor 1992)
c. Guess the **color** of my eyes.

We propose that it is functional nouns (in the sense of Vinker & Jensen 2002) that allow for CQ interpretation, i.e. nouns whose interpretation depends on an additional argument. The nouns in (44) are functional nouns denoting people: a person is not a murderer by virtue of some properties inherent to the person himself; rather, that person must be a murderer of someone. The nouns in (45) are functional nouns denoting locations. In (46) the output of the function is a certain number. The nouns in (47) are other functional nouns.

In (48), the nouns themselves are not functional, but the whole phrase is. For example, while the noun *person* is not functional, the nominal *the person who won the last Pulitzer Prize* in (48b) is.

(48) a. Tell me your **favorite movie**.
b. Tell me **the person who won the last Pulitzer Prize**.
c. Tell me **the candy Jill wants to buy**.
d. John can't remember **the wine she likes**.
e. Tell me **the largest city in Italy**.

However, this cannot be the whole story. In particular, the nominal in (40b) *the city you live in* is also functional – it is a function from you to the place you live in. While we believe that the generalization that only functional nominals are possible CQs is on the right track, a more fine grained notion of functional is clearly needed. We leave this issue here – see Nathan (forthcoming) for further development of this idea.

4. A Note on the Status of the Post-copular Phrase

Although this chapter focuses on the status of the pre-copular phrase as a question, it is important to point out in this context the existence of "anti-connectivity" effects in specificational sentences (originally discussed in Sharvit 1999) which are relevant for the status of the post-copular phrase. Anti-connectivity involves cases where the "connected" simple sentence exhibits the opposite pattern from the specificational pseudocleft with respect to the range of syntactic and semantic phenomena known as connectivity effects.

Anti-connectivity effects are exemplified in (49-52). (49a-b) are examples of Principle A anti-connectivity, (50) demonstrates Principle B anti-connectivity, and (51) is an example of Principle C anti-connectivity (in Italian). (52) exemplifies NPI anti-connectivity.

(49) *Principle A anti-connectivity*
a. The person every professor hopes his wife likes is himself. (Sharvit 1999)
→ *Every professor hopes his wife likes himself.
b. What John thinks that Mary likes is himself (Schlenker 2003)
→ *John thinks that Mary likes himself

(50) *Principle B anti-connectivity*
??The person every professor hopes his wife likes is him. (Sharvit 1999)
→ Every professor hopes his wife likes him.

(51) ITALIAN

Principle C anti-connectivity (in Italian) (Cecchetto 2001)
*Chi lo_i vide è la sorella di $Gianni_i$
Who him saw is the sister of Gianni
'The person who saw him_i is $Gianni_i$'s sister.'
→ La sorella di $Gianni_i$ lo_i vide
the sister of Gianni him saw
'$Gianni_i$'s sister saw him_i.'

(52) *NPI anti-connectivity* (den Dikken et al. 2000)
?What WASn't sitting on the shelf was a book that said anything sensible about X
→ A book that said anything sensible about X wasn't sitting on the shelf.

The existence of anti-connectivity effects shows that connectivity effects cannot always be accounted for via ellipsis of a full IP in the post-copular position, as ellipsis would yield the wrong result in these cases. Anti-connectivity effects, then, argue against any 'Question plus Deletion' analysis, such as den Dikken et al. (2000) and Schlenker (2003). Note that these effects also argue against any other analysis of connectivity which uses a full IP.

These effects, however, do not argue against <u>all</u> versions of the question-answer approach. In particular, as pointed out by Schlenker (2003), the same anti-connectivity effects are found in question-short answer pairs. (53-55) show for question-answers the anti-connectivity effects we saw in pseudoclefts in (49), (51) and (52) respectively.

(53) *Principle A anti-connectivity* (Schlenker 2003)
What does John think that Mary likes? – Himself.

(54) ITALIAN

Principle C anti-connectivity (in Italian) (Cecchetto 2001)
*Chi lo_i vide? – La sorella di $Gianni_i$
Who him saw The sister of Gianni
'*Who saw him_i? – $Gianni_i$'s sister.'

(55) *NPI anti-connectivity* (Schlenker 2003)
?What WASn't sitting on the shelf? – A book that said anything about X.

The parallelism between questions and specificational sentences may be used to argue for a new version of the question-answer approach where a specificational sentence is a pair of question-short answer. Note, however, that the parallelism by itself does not constitute an account of (anti-)connectivity effects – a theory that accounts for these effects without using the notion of c-command must be developed in order to make such an analysis of specificational sentences a real option.

Crucially, the arguments presented in this chapter against analyzing the pre-copular phrase in a specificational sentence as a question, whether syntactically or semantically, are relevant for any possible question-short answer analysis as well. That is, while this version of the question-answer approach is more attractive from the perspective of accounting for connectivity (and anti-connectivity), it faces the same problems concerning the motivation for (or, rather, the lack of motivation for) analyzing the pre-copular phrase as a question.

5. CONCLUSIONS

This chapter has provided crosslinguistic evidence against the question-answer approach to connectivity in specificational sentences. We focused on the status of the pre-copular phrase as a question, as it constitutes the motivation for the question-answer approach. In addition, since being an answer concerns the discourse status of an indicative sentence, it is easier to examine the pre-copular phrase as a question which is expected to have certain syntactic or semantic properties. If the pre-copular phrase in a specificational sentence is not a question, then the post-copular phrase cannot be an answer and we

cannot explain connectivity by relating it to parallel effects found in question-answer pairs (independent of whether we use short or full answers, as was shown in section 4).

We presented two distributional arguments based on new data from five languages – English, Macedonian, Hebrew, Wolof and Hungarian – against analyzing the *wh*-clause in a specificational pseudocleft as an embedded *wh*-interrogative, contra den Dikken *et al.* (2000). We then presented three arguments showing that the pre-copular phrase in a specificational sentence is also not interpreted like a question, i.e. it is not a CQ, contra Schlenker's (2003) proposal. These arguments concerned distributional differences between the pre-copular position of specificational sentences and canonical CQ environments, but also targeted the interpretation of phrases occurring in the pre-copular position of specificational sentences directly. Again, the data was derived from the same five genetically unrelated languages – English, Macedonian, Hebrew, Wolof and Hungarian.

Having seen evidence that the pre-copular phrase in a specificational sentence is not a question, we conclude that the question-answer approach is not the right approach for specificational sentences. It is important to note that this does not rule out the existence of question-answer equations in natural language, but merely the analysis of specificational sentences as such. If specificational sentences are not question-answer pairs, connectivity cannot be accounted for by exploiting the parallelism with similar effects in question-answer pairs.

The next chapter presents an alternative approach to connectivity, in which specificational sentences are taken to be copular sentences as we see on the surface and connectivity effects are derived as a by-product of equation. We will use this approach,

originally developed by Jacobson (1994) and Sharvit (1999) for English connectivity, to develop a detailed analysis of the complex connectivity pattern in Hebrew.

Chapter 3

A Direct Compositional Analysis of Connectivity in Hebrew*

In this chapter I argue for an alternative to the question-answer approach to connectivity we saw in the previous chapter. In particular, I adopt the direct compositional approach, originally developed in Jacobson (1994) and Sharvit (1999), which follows the spirit of Higgins' (1973) null hypothesis in assuming that "no specific pseudocleft transformation is involved in their derivation" (Higgins, p. 13). Instead, each connectivity effect is reanalyzed such that its analysis does not rely on the syntactic configuration of c-command. That is, this approach takes the existence of connectivity effects as evidence that the notion of c-command is not relevant to the licensing of these different syntactic and semantic phenomena. The goal is to develop alternative analyses that do not make use of c-command for each of the phenomena. This analysis should be able to account for all cases, including those that are perceived as "connectivity". Such analyses have been developed for a number of connectivity effects: Jacobson (1994) presents an analysis of Principle A connectivity and Bound Variable (BV) connectivity; Sharvit (1999) presents an analysis of opacity connectivity, NPI connectivity and sketches an analysis of Principle B connectivity; Cecchetto (2000, 2001) presents an analysis of quantifier scope connectivity.

This chapter extends the direct compositional approach to Hebrew, which exhibits a complex connectivity pattern. We will see that Hebrew has two types of specificational pseudoclefts that each exhibit a different subset of connectivity effects. The direct

---

* The analysis presented in this chapter has been adopted with minor changes from Heller (2002). Some generalizations in that paper are drawn from Heller (1999).

compositional account that deals with each connectivity effect separately can naturally account for such a pattern. In addition, Hebrew has two connectivity effects that are not found in English – agreement connectivity and Accusative marking by *et* – for which I develop a direct compositional account. In order to have a manageable data set, I will concentrate on pseudoclefts, i.e. on specificational sentences in which pre-copular phrase is a FR. However, the analysis is relevant for non-pseudoclefted specificational sentences as well.

The chapter is organized as follows. Section 1 previews the connectivity pattern in Hebrew. Section 2 takes a step back and examines the range of pseudoclefts in Hebrew. We will see that Hebrew distinguishes predicational and specificational pseudoclefts in the choice of the copula, and that both types of pseudoclefts presented in section 1 behave like English specificational pseudoclefts on various structural tests. This section reproduces arguments that were first presented in Heller (1999). In section 3, which is the main section of the chapter, we turn to the analysis of connectivity. The direct compositional analyses developed for English will be extended to account for the complex Hebrew pattern. This section follows Heller (2002) with some minor changes. Finally, section 4 reevaluates the analyses of connectivity we saw in previous chapters in light of the Hebrew data, focusing on the question-answer approach. We will see that no approach other than the direct compositional approach can deal with the complex pattern of connectivity in Hebrew. Throughout the chapter, all the examples are in English or in Hebrew unless otherwise noted.

1. The Connectivity Pattern in Hebrew

Specificational sentences in Hebrew exhibit connectivity effects similar to the connectivity effects we saw in the previous chapter for English. One such effect is the binding of a reflexive across the copula. (1a) exemplifies Principle A connectivity and (1b) is the corresponding simple sentence.

(1) a. [ma še-**dan** haya _] ze mesukan le-**acmo**
what that-Dan was Z dangerous to-himself
'What Dan was was dangerous to himself.'

b. **dan** haya mesukan le-**acmo**
Dan was dangerous to-himself
'Dan was dangerous to himself.'

Note that the copula in this example is what is known as the impersonal pronominal copula (glossed as *Z*) – the nature of the copula will be discussed in detail in section 2.

In addition, Hebrew also has some connectivity effects that are not found in English. One such effect is Accusative marking by *et*, which obligatorily marks all definite direct objects. (2) demonstrates that *et* obligatorily marks a (definite) direct object, but is banned from subject position.

(2) a. kaninu *(et) ha-xulca ha-kxula ba-šuk
we-bought Acc the-shirt the-blue in-the-market
'We bought the blue shirt in the market.'

b. (*et) ha-xulca ha-kxula nikre'a ba-tiyul
Acc the-shirt the-blue tore in-the-trip
'The blue shirt tore during the trip.'

In the pseudoclefted counterparts in (3), the post-copular definite NP must be marked by *et* when the gap in the FR is in object position, and must not be marked by *et* when the gap is in subject position.

(3) a. [ma še-kaninu ba-šuk _] ze *(et) ha-xulca ha-kxula
what that-we-bought in-the-market Z Acc the-shirt the-blue
'What we bought in the market is the blue shirt.'

b. [ma še _ nikra ba-tiyul] ze (*et) ha-xulca ha-kxula
what that tore in-the-trip Z Acc the-shirt the-blue
'What tore during the trip was the blue shirt.'

Unlike in English, in Hebrew not all specificational pseudoclefts exhibit connectivity. In particular, Hebrew has two types of specificational pseudoclefts: those where the (impersonal) pronominal copula agrees with the post-copular phrase, and those where it has a fixed "neutral" form (which is identical in form to the masculine singular agreeing copula). Interestingly, the form of the copula affects the availability of connectivity effects. The contrast in (4) demonstrates that the connectivity effect of *et* marking shows up on the (feminine) post-copular NP only when the copula is neutral (all nouns in Hebrew are marked for gender).

(4) a. [ma še-kaninu ba-šuk _] ze *(et) ha-xulca ha-kxula
what that-we-bought in-the-market Z(n) Acc the-shirt(f) the-blue(f)
b. [ma še-kaninu ba-šuk _] zot (*et) ha-xulca ha-kxula
what that-we-bought in-the-market Z(f) Acc the-shirt(f) the-blue(f)
both: 'What we bought in the market was the blue shirt.'

In (4a), repeated from (3a), the copula is in its neutral form, and *et* is obligatorily present in accordance with the position of the gap in the FR (*ze* cannot be the agreeing masculine form here, because the post-copular NP is feminine). In (4b) the copula agrees with the post-copular NP, and *et* is obligatorily absent. In sum, the pseudocleft with the neutral copula is "connected" with respect to *et* marking, while the one with the agreeing copula is not.

A second connectivity effect that shows up only when the copula is in its neutral form is opacity, illustrated by the contrast in (5). Again, since the masculine agreeing copula cannot be distinguished from the neutral copula, the post-copular phrase (and hence the agreeing copula) is feminine.

(5) a. [ma še-dan mexapes _] ze jirafa vruda **de dicto & de re**
what that-Dan seeks Z(n) giraffe(f) pink(f)
'What Dan seeks is a pink giraffe.'

b. [ma še-dan mexapes _] zot jirafa vruda **only de re**
what that-Dan seeks Z(f) giraffe(f) pink(f)
de re of 'What Dan seeks is a pink giraffe.'

(5a) is an instance of opacity connectivity as it is ambiguous between a *de dicto* reading (where the existence of pink giraffes is not entailed) and a *de re* reading (where the existence of pink giraffes is entailed), even though the NP is not in the semantic scope of the intensional verb. This ambiguity is not found in (5b) where the copula agrees with the post-copular feminine NP: the sentence only has the *de re* reading, as expected from the fact that the NP is not in the scope of the intensional verb.

But not all connectivity effects are absent from pseudoclefts with the agreeing copula. One connectivity effect that is found with both types of copulas is Bound Variable connectivity, illustrated in (6) (special thanks to Yael Sharvit for drawing my attention to this example).

(6) mi še-**[kol gever]$_i$** ohev ze /zot ima šel-**o$_i$**
who that-every man loves Z(n)/Z(f) mother of-his
'The person every man loves is his mother.'

Here, both the neutral copula version and the agreeing copula version allow for the bound reading, i.e. the sentence is true if every man loves his own mother. That is, unlike the connectivity effects of opacity and *et* marking that are only found in the neutral copula pseudoclefts, BV connectivity is found with both types of copulas.

The emerging descriptive generalization is that Hebrew has two identifiable subsets of specificational pseudoclefts that behave differently with respect to connectivity. In particular, when the copula is in its neutral form, the pseudocleft exhibits connectivity effects (we will see two exceptions in section 3), and when the copula is in its agreeing

form, only some connectivity effects are found. This "mixed-bag" pattern shows that there is no one-to-one correlation between specificational pseudoclefts and connectivity effects, thereby suggesting that connectivity is not a defining characteristic of specificational sentences. This implies that connectivity should not be derived by a single grammatical operation applied to all specificational sentences, but rather be accounted for independently of each other. This is the line I take in section 3. But, first, let us take a closer look at the range of pseudoclefts in Hebrew.

2. THREE PSEUDOCLEFTS IN HEBREW

2.1. *Two pronominal copulas and two pseudoclefts*

In this section we take a step back and look at the full picture of Hebrew pseudoclefts. Recall that in English there are two types of pseudoclefts, predicational and specificational, as exemplified by the ambiguous pseudocleft in (7), repeated from chapter 1 example (5).

(7) What John is is important.

Against the background of this ambiguity, let us consider Hebrew pseudoclefts. It is known since the work of Berman & Grosu (1976) and Doron (1983) that Hebrew lacks a present-tense verbal copula, and instead it employs a personal or impersonal pronoun in this role. Although these pronominal copulas are historically derived from subject pronouns, they have been demonstrated to be copular elements and not subject pronouns (for more on the nature and distribution of pronominal copulas see the above references and also Rapoport 1987, Rothstein 1995, 2000, Greenberg 1994, 1998, Sichel 1997 and Heller 1999). Adapting the terminology of Doron (1983), I refer to the personal

pronominal copulas as pronH and to the impersonal pronominal copula as pronZ. These two pronominal copulas yield two versions of the ambiguous English pseudocleft.

(8) a. ma še-dan haya hu mo'il la-xevra **pronH**
what that-Dan was H helpful to-the-society **only predicational**
'What Dan was is helpful to society.'

b. ma še-dan haya ze mo'il la-xevra **pronZ**
what that-Dan was Z helpful to-the-society **only specificational**
'What Dan was was helpful to society.'

The meaning contrast in the minimal pair in (8) corresponds with the two types of pseudoclefts in English: the pronH version (8a) has the predicational reading and the pronZ version (8b) has the specificational reading[1]. I thus propose that the two pronominal copulas induce the two types of pseudoclefts, i.e. that the distinction between predicational and specificational pseudoclefts is overtly marked in Hebrew in the choice of the copula.

One piece of evidence for this hypothesis comes from the pseudoclefts we saw in section 1 which were all pronZ pseudoclefts (the Hebrew pseudoclefts in the first two chapters were also pronZ pseudoclefts). These pseudoclefts were shown to exhibit connectivity on a par with specificational pseudoclefts in English. If pronH pseudoclefts are indeed predicational, they are expected not to exhibit connectivity. This prediction is borne out, as exemplified in (9) for the connectivity effect of licensing a reflexive across the copula.

[1] Unlike their English translations, Hebrew specificational pseudoclefts do not exhibit tense harmony, i.e. the form of the copula in Hebrew is not affected by the tense of the verb inside the FR. This is due to the different nature of their copulas: while the English copula is a tensed verb, the pronominal copula in Hebrew, like the so-called "present-tense" verbs, is actually unspecified for tense (Doron 1983).

(9) a. [ma še-**dan** haya _] ze mesukan le-**acmo**
what that-Dan was Z dangerous to-himself
'What Dan was was dangerous to himself.'

b. *[ma še-**dan** haya _] hu mesukan le-**acmo**
what that-Dan was H dangerous to-himself
'*What Dan was is dangerous to himself.'

The pronZ version (9a), repeated from (1b), licenses the bound reflexive reading, i.e. it is "connected" like English specificational pseudoclefts. The ungrammaticality of (9b) is due to the fact that the reflexive in the post-copular phrase is not licensed by the antecedent inside the FR, i.e. the pronH pseudocleft is not "connected", parallel to predicational pseudoclefts in English. This contrast in the availability of connectivity effects provides preliminary support to the correlation between the type of the copula and the type of the pseudocleft.

Further support for this correlation comes from structural tests that Higgins (1973) introduces for distinguishing predicational and specificational pseudoclefts. The rest of this section presents the original English tests along with their Hebrew equivalents. Note that the pronZ examples used throughout this section allow only for neutral pronZ, as the post-copular phrase is a predicate (see section 3 for analysis). Section 2.2 will present the same tests for agreeing pronZ pseudoclefts.

Test 1: Coordination of Predicates. Since predicates can only be coordinated with other predicates, coordinating an ambiguous phrase with an unambiguous predicate yields an unambiguous predicational pseudocleft. Thus, adding a verbal predicate to the ambiguous pseudocleft in (10a) yields the unambiguously predicational pseudocleft in (10b).

(10) a. What I'm pointing at is a cat.
b. What I'm pointing at is a cat and is called Jemima. (Higgins, p.213)

(11), the Hebrew equivalent of (10b) is possible with pronH but not with pronZ, showing that a pronH pseudocleft is predicational and that a pronZ pseudocleft is specificational.

(11) ma še-ani macbi'a alav hu/*ze xatul ve-mexune garfild
what that-I point on-it H/*Z cat and-called Garfield
'What I'm pointing at is a cat and is called Garfield.'

Note that unlike in English where the copula is repeated with the verbal predicate, the verbal predicate in Hebrew occurs without the copula. This difference is not a result of coordination, but rather a characteristic of passive predicates in Hebrew which never occur with a copula, e.g. *ha-xatul (*hu/ze) mexune garfild* 'The cat is called Garfield'.

Test 2: Extraction out of the Post-Copular Phrase. Higgins (p. 308) points out that extraction out of the post-copular phrase is possible (although not perfect) in predicational pseudoclefts, as in (12), but not in specificational ones, as in (13).

(12) a. They said that what Mary was looking at appeared to be [a picture of a kangaroo].
b.?What did they say that what Mary was looking at appeared to be [a picture of _]?

(13) a. They said that what Mary was going to do was [give the dog to John].
b. *Who did they say that what Mary was going to do was [give the dog to _] ?

In Hebrew, extraction is possible only out of the pronH pseudocleft in (14b), but not out of the pronZ pseudoclefts in (15b), providing further support to the correlation between pronH and predicational pseudoclefts and between pronZ and specificational pseudoclefts.

(14) a. hem xošvim še-ma še-dan bana hu/*ze [mo'il la-xevra]
they think that-what that-Dan built H/*Z helpful to-the-society
'They think that what Dan built is helpful to society.'

b. lemi hem xošvim še-ma še-dan bana hu [mo'il _] ?
to-who they think that-what that-Dan built H helpful
'To whom do they think that what Dan built is helpful?'

(15) a. hem xošvim še-ma še-dan haya *hu/ze [nexmad la-orxim]
they think that-what that-Dan was *H/Z nice to-the-guests
'They think that what Dan was was nice to the guests.'

b. *le-mi hem xošvim še-ma še-dan haya ze [nexmad _] ?
to-who they think that-what that-Dan was Z nice
'*To whom do they think that what Dan was was nice?'

Test 3: Deletion of the Post-Copular Phrase. An additional distinction between predicational and specificational pseudoclefts shows up if we try to delete the post-copular phrase in an appropriate context. (16a) is a coordination of two ambiguous pseudoclefts: on the predicational reading, John and Mary's jobs are important and interesting, and on the specificational reading, it is John and Mary themselves who are important and interesting. Higgins (p. 302) shows that deleting the post-copular phrase from the second pseudocleft causes the specificational reading to disappear, i.e. such deletion is only possible in a predicational pseudocleft.

(16) a. What John is is important and what Mary is is interesting.
b. What John is is important and what Mary is is too.

In Hebrew, adding an elided pseudocleft to the pseudocleft in (8) renders only the pronH version grammatical, supporting the hypothesis that it is predicational, and that pronZ pseudoclefts are specificational.

(17) a. ma še-dan haya hu mo'il la-xevra ve-ma še-ron haya gam
what that-Dan was H helpful to-the-society and-what that-Ron was also
'What Dan was is helpful to society and what Ron was is too.'

b. *ma še-dan haya ze mo'il la-xevra ve-ma še-ron haya gam[2]
what that-Dan was Z helpful to-the-society and-what that-Ron was also
'*What Dan was was helpful to society and what Ron was was too.'

---

[2] Note that this string of words has a grammatical reading as left dislocation (which differs from the copular sentence in intonation): the FR serves as the dislocated element, *ze* is the subject pronoun (not a pronominal copula!) referring back to the FR, and *mo'il la-xevra* 'helpful to society' is a predicate, predicated over the subject pronoun. Since this is a predicational sentence, the elided reading is expected to be grammatical.

Note that deleting the post-copular phrase in Hebrew requires the deletion of the copula as well. Therefore, this test may be a special case of the next test which deals with the deletion of the copula alone. In any case, both tests yield the same result.

Test 4: Deletion of the Copula. Higgins (p. 305) shows that the second copula in a coordination of pseudoclefts can only be omitted in a coordination of two predicational pseudoclefts. So while the coordination in (18a) is ambiguous, deleting the second copula renders (18b) unambiguously predicational, i.e. it can only mean that John and Mary's jobs are important and interesting, but not that John and Mary themselves are important and interesting.

(18) a. What John is is important and what Mary is is interesting.
b. What John is is important and what Mary is ... interesting.

An analogous situation is found in Hebrew. Adding a second pseudocleft from which the copula has been omitted is possible to a pronH pseudocleft but not to a pronZ pseudocleft.

(19) a. ma še-dan haya hu/ze mo'il la-xevra ve-ma še-ron haya
what that-Dan was H/Z helpful to-the-society and-what that-Ron was
hu/ze mezik la-xevra.
H/Z harmful to-the-society
'What Dan was is/was helpful to society and what Ron was is/was harmful to society.'

b. ma še-dan haya hu/*ze mo'il la-xevra ve-ma še-ron haya …
what that-Dan was H/*Z helpful to-the-society and-what that-Ron was …
mezik la-xevra
harmful to-the-society
'What Dan was is/*was helpful to society and what Ron was … harmful to society.'

Test 5: Referentiality of the FR. The last test distinguishing predicational and specificational pseudoclefts is concerned with the referentiality status of the pre-copular FR. According to Higgins (1973: 264), the pre-copular position in a predicational pseudocleft hosts a referential expression, while that of a specificational pseudocleft hosts

a non-referential one. Higgins claims that the negated FR in (20) is non referential in the sense that it cannot pick out an object in the world. As a result, (20) is unambiguously specificational[3].

(20) What I don't like about John is his tie. (Higgins, p.214)

We will see in chapter 5 that when the FR in (20) is combined with a different post-copular phrase it can occur in the pre-copular position of a predicational sentence. However, what is important for our purposes here is what copulas are allowed for this sentence in Hebrew. (21) shows that the pronZ version in (21b) is grammatical and the pronH version in (21a) is ungrammatical.

(21) a. *ma še-ani lo ohevet ecel dan hu (et) ha-aniva shelo
what that-I not love about Dan H Acc the-tie his
b. ma še-ani lo ohevet ecel dan ze *(et) ha-aniva shelo
what that-I not love about Dan Z Acc the-tie his
'What I don't like about Dan is his tie.'

Note that in the pronZ version the post-copular phrase must be marked by *et* – this is a case of *et* marking connectivity which we will discuss in section 3.3. Crucially, the pronH version is ungrammatical independent of the presence or absence of the Accusative marker *et*, i.e. the pronH version is not available when the predicational reading is expected to be absent.

The fact that throughout these tests pronH is blocked when the predicational reading is absent and pronZ occurs when the specificational reading is available shows that the hypothesis that the two copulas – pronH and pronZ – yield the two types of pseudoclefts

[3] The original reason Higgins discusses this sentence is not to distinguish predicational and specificational sentences but rather as evidence against deriving specificational pseudoclefts from the "corresponding" simple sentences. In particular, (20) cannot be derived from the corresponding simple sentence because this sentence is ungrammatical, as in (i)
(i) *I don't like his tie about John.

– predicational and specificational is correct. The fact that the choice of the copula determines the type of the pseudocleft in Hebrew contributes morphological evidence to the long standing distinction between predicational and specificational pseudoclefts. This overt distinction enables us investigate the two types of pseudoclefts independently and could be used to evaluate the status of copular sentences that are not easily classified as belonging to either type. We will use this overt distinction in chapter 5.

2.2. *A third pseudocleft*

At this point we have introduced three types of pseudoclefts in Hebrew: predicational pronH pseudoclefts (section 2.1), specificational neutral pronZ pseudoclefts that exhibit connectivity effects (sections 1 and 2.1), and agreeing pronZ pseudoclefts that show only a limited range of connectivity effects (section 1). All are illustrated in (22).

(22) ma še-dekart maca hu /ze /zot hoxaxa le-kiyum ha-el
what that-Descartes found H(m)/Z(n)/Z(f) proof(f) to-existence the-god
'What Descartes found is/was a proof of god's existence.'

The type of agreeing pronZ pseudoclefts is yet to be determined. The fact that this copula is historically derived from the impersonal pronoun like neutral pronZ suggests that the pseudocleft it induces is specificational on a par with the one induced by neutral pronZ. On the other hand, if connectivity is a defining characteristic of specificational sentences, the absence of a number of connectivity effects from agreeing pronZ pseudoclefts casts doubt on their being specificational. In order to better understand the nature of these pseudoclefts, we can use the structural tests (other than connectivity) that we saw in the previous section. (23) illustrates tests 1-4 which are conducted for (22). (24) illustrates test 5.

(23) a. *Coordination of Predicates.*
ma še-dekart maca hu /*ze /*zot hoxaxa le-kiyum ha-el ve-
what that-Descartes found H(m)/Z(n)/Z(f) proof(f) to-existence the-god and-
mofi'a be-kol sifrey ha-filosofya
appears in-all books the-philosophy
'What Descartes found is a proof of god's existence and is found in all philosophy books.'

b. *Extraction out of the Post-Copular Phrase.*
le-kiyumo šel mi ma še-dekart maca hu /??ze /*zot [hoxaxa _]?
to-existence of who what that-Descartes found H(m)/??Z(n)/*Z(f) proof
'Of whose existence is what Descartes found a proof?'

c. *Deletion of the Post-Copular Phrase.*
ma še-dekart maca hu /*ze /*zot hoxaxa le-kiyum ha-el
what that-Descartes found H(m)/Z(n)/Z(f) proof(f) to-existence the-god
ve-ma še-špinoza maca gam
and-what that-Spinoza found also
'What Descartes found is a proof of god's existence and what Spinoza found is too.'

d. *Deletion of the Copula.*
ma še-dekart maca hu /*ze /*zot hoxaxa le-kiyum ha-el
what that-Descartes found H(m)/Z(n) /Z(f) proof(f) to-existence the-god
ve-ma še-špinoza maca … hoxaxa le-he'ader ha-el
and-what that-Spinoza found proof(f) to-absence the-god
'What Descartes found is a proof of god's existence and what Spinoza found … a proof of the absence of god.'

(24) *Referentiality of the FR Subject.*
a. *ma še-ani lo ohevet ecel dan hu (et) ha-aniva shelo
what that-I not love about Dan H Acc the-tie his
b. ma še-ani lo ohevet ecel dan ze *(et) ha-aniva shelo
what that-I not love about Dan Z Acc the-tie his
c. ma še-ani lo ohevet ecel dan zot (*et) ha-aniva shelo
what that-I not love about Dan Z Acc the-tie his
all: 'What I don't like about Dan is his tie.'

Agreeing pronZ patterns with neutral pronZ on all five tests. Note that the agreeing and the neutral pronZ versions in (24) differ in their connectivity behavior – we come back to that in Section 3.3. This suggests that agreeing pronZ pseudoclefts are not predicational, but rather specificational like neutral pronZ pseudoclefts.

If we classify agreeing pronZ pseudoclefts as specificational based on the form of the copula and their behavior on the tests, we would have to explain why connectivity effects that are known to characterize specificational pseudoclefts are absent in part from agreeing pronZ pseudoclefts. In this context, the connectivity effects that are found in these pseudoclefts also need to be accounted for. This is the line I pursue here.

It should be pointed out here that along with predicational and specificational pseudoclefts, Higgins identifies a third type of pseudoclefts – identificational pseudoclefts. Higgins does not discuss identificational pseudoclefts in the context of the other structural tests, including connectivity, but we will see in chapter 5 (section 3.1) that non-pseudoclefted identificational sentences pattern with predicational sentences on most of Higgins' tests (pseudoclefts cannot be tested because all the pseudoclefts that are claimed by Higgins to be identificational are ambiguous). Thus, based on their behavior on the structural tests, it is inappropriate to classifying agreeing pronZ pseudoclefts as identificational. Moreover, if we adopt such classification, we would have to come up with an explanation for why these pseudoclefts exhibit some connectivity effects.

Note that independent of the label we choose to assign to agreeing pronZ pseudoclefts, we have a situation where there is no one-to-one correlation between connectivity and specificational pseudoclefts. Such a paradigm cannot be accounted for if we take connectivity effects to be a result of any one grammatical operation, e.g. any form of reconstruction, copying or ellipsis. For convenience, let us refer to this option of deriving connectivity using the general term "operation C(onnectivity)" and postpone the discussion of the specific analyses until section 4. What we need in order to account for this paradigm is an approach where the different connectivity effects are independent of

each other. The direct compositional approach of Jacobson (1994) and Sharvit (1999) (also known as the "semantic theory") takes each connectivity effect to be a by-product of certain semantic combinatorics, and the denotations of the different elements are independently motivated for other environments. Under this approach, it is not surprising that certain restrictions on the pseudocleft may result in a situation where some connectivity effects are available while others are blocked. The next section presents an analysis of the presence and absence of connectivity effects in both types of pronZ pseudoclefts.

3. The Direct Compositional Analysis of Connectivity

Under the direct compositional approach to connectivity (Jacobson 1994, Sharvit 1999), connectivity effects are a by-product of identity semantics at the sentence level – the way in which identity is composed does not affect the derivation of connectivity[4]. We saw in chapter 1 that in addition to deriving identity directly, identity can be composed directly or as a special case of predication or inverse predication. In Heller (1999, chapter 3) I have argued that specificational (i.e. pronZ) pseudoclefts in Hebrew are identity statements, i.e. identity is composed directly, and I will adopt this analysis here. In chapter 5 (section 2) I discuss and reject the inverse predication analysis. Nonetheless, it should be pointed out that the analysis presented here only depends on specificational sentences expressing identity at the sentence level.

To analyze the types of pronZ pseudoclefts, I propose that they differ in the arguments they equate. (25) illustrates that neutral pronZ can occur with any type of post-

---

[4] Indeed, Jacobson and Sharvit differ in their analyses: Jacobson (in her 1995 paper) adopt the inverse predication analysis while Sharvit does not take a stand on the matter.

copular phrase, while agreeing pronZ is restricted to some NPs. Note that this does not simply follow from the requirement that agreeing pronZ would agree in gender (and number) with the post-copular phrase, as both VPs and APs in Hebrew are marked for gender.

(25) a. ma še-rut xašva ze/*zot **[**še-yered maxar gešem**]** **CP**
what that-Ruth thought Z(n)/Z(f) that-go-down tomorrow rain
'What Ruth thought was that it would rain tomorrow.'

b. ma še-rut asta ze/*zot **[**halxa ha-bayta**]** **VP**
what that-Ruth did(f) Z(n)/Z(f) went(f) to-the-house
'What Ruth did was went home.'

c. ma še-rut hayta ze /*zot **[**mo'ila la-xevra**]** **AP**
what that-Ruth was(f) Z(n)/Z(f) helpful(f) to-the-society
'What Ruth was was helpful to society.'

d. ma še-rut crixa ze /*zot **[**beyn xameš le-ševa beycim**]** **NP**[5]
what that-Ruth needs Z(n)/*Z(f) between five to-seven eggs(f)
'What Ruth needs is between five and seven eggs.'

e. ma še-dekart maca ze /zot **[**hoxaxa le-kiyum ha-el**]** **NP**
what that-Descartes found Z(n)/Z(f) proof(f) to-existence the-god
'What Descartes found was a proof of god's existence.'

---

[5] Not all quantificational NPs are possible with neutral pronZ (agreeing pronZ is impossible here).

(i) a. ma še-rut kanta ze štey xovrot be-balšanut
what that-rut bought Z(n) two booklets(f) in-linguistics
'What Ruth bought was two booklets on linguistics.'
b. ??ma še-rut kanta ze le-faxot šaloš xovrot be-balšanut
what that-Ruth bought Z(n) at least three booklets(f) in-linguistics
'What Ruth bought was at least three booklets on linguistics.'
c. *ma še-rut kar'a ze kol katava al xomski
what that-Ruth read Z(n) every article about Chomsky
'What Ruth read was every article about Chomsky.'

While the data in (i) seem to fit Heycock & Kroch's (1999) generalization about English that only (plural) individual denoting quantificational NPs can serve as the post-copular element in a specificational pseudocleft, (25d) shows that this generalization is not true for Hebrew. Understanding the distribution of post-copular quantificational NPs requires a theory of quantifiers in Hebrew which is currently unavailable. However, in this context it is interesting to point out that Hebrew does not exhibit NPI connectivity – this may be related to the limitations on which quantifiers can occur in the post-copular position of specificational pseudoclefts.

(ii) (dan maca kol miney katavot, aval) *ma Se-hu lo maca ze af katava relevantit
Dan found all kinds articles but what that-he not found Z(n) any article(f) relevant(f)
"Dan found all kinds of articles, but what he didn't find was any relevant article."

Our initial generalization is that neutral pronZ is cross-categorial, as suggested by Partee (1986) for the English *be*, whereas agreeing pronZ is restricted to entity-denoting NPs. In many models, e.g. Chierchia (1984), the domain of entities[6] (type e) is sorted such that it also contains entities that do not denote objects, but rather properties and propositions (and potentially other sorts as well). Interestingly, agreeing pronZ is restricted to a subset of entities. This is illustrated by (26a), repeated from (25e), which contrasts with the post-copular property-like entity in (26b), the proposition-like entity in (26c), and the event denoting nominal in (26d) which are all impossible with agreeing pronZ.

(26) a. ma še-dekart maca ze /zot hoxaxa le-kiyum ha-el
what that-Descartes found Z(n)/Z(f) proof(f) to-existence the-god
'What Descartes found was a proof of god's existence.'

b. ma še-rut hayta ze /*zot madrixat tiyulim
what that-Ruth was(f) Z(n)/Z(f) guide(f) tours
'What Ruth was was a tour guide.'

c. ma še-dan daxa ze /*zot te'ana bidvar šxitut ha-ma'arexet
what that-Dan rejected Z(n)/Z(f) claim(f) about corruption the-system
'What Dan rejected was a claim about the corruption of the system.'

d. ma še-ha-xazay cafa ze/*zot acirat gšamim
what that-the-meteorologist forecasted Z(n)/Z(f) stopping(f) rains
ba-xoref ha-karov
in-the-winter the-close
'What the meteorologist forecasted was stopping of the rain for next winter.'

The data in (26) shows that pronZ occurs with just a subset of type e entities. In addition, we will see in section 3.2 that agreeing pronZ can occur with functions of type <e,e> whose range is the same subset of entities. We are therefore led to conclude that the

[6] I use Montague's original term 'entities' instead of the common term 'individuals' in order to be clear that we are dealing with an ontological concept rather than with a natural class of things in the world. Individuals will only be used to refer to people.

formal restriction on the arguments equated by agreeing pronZ does not refer to all of the entities in the logical type e, but rather to a subset of this type. This subset has to exclude abstract entities that correlate with predicates, propositions and events, but it must at the same time include abstract nouns like *hoxaxa* 'proof' (26a), in addition to concrete nouns like *jirafa* 'giraffe' (5) and *xulca* 'shirt' (4). Since defining this class is not straightforward, I will not provide such a definition here. I will refer to this subset as "real world" entities because these are the denotations that are intuitively entities in the world, in contrast with properties, propositions and events which intuitively are not entities. In any case, the fact that the grammar is sensitive to (some notion of) "real world" entities indicates the importance of sorts in the domain of type e.

The rest of this section accounts for the presence and absence of each of the connectivity effects in neutral and agreeing pronZ pseudoclefts. The pattern will be a result of the difference in semantic type between the arguments of the two pseudoclefts combined with semantic denotations adopted in part from Jacobson (1994) and Sharvit (1999). I show that the restriction on agreeing pronZ to equating "real world" entities blocks many connectivity effects, while the cross-categorial nature of neutral pronZ will allow for almost all connectivity effects to arise. Throughout the section, I will point out why this mixed-bag pattern cannot be accounted for by any single operation that posits c-command relation between the phrases around the copula at some level of representation (*Operation C*).

3.1. *Opacity connectivity*

We have already seen that neutral pronZ pseudoclefts are "connected" with respect to opacity – the post-copular phrase is ambiguous between a *de dicto* and a *de re* reading

even though it is not in the syntactic, and hence the semantic, scope of the intensional verb which is located inside the FR. This contrasts with agreeing pronZ pseudoclefts where only the de re reading is available. Consider (27), repeated from (5).

(27) a. ma še-dan mexapes ze jirafa vruda **de dicto & de re**
what that-Dan seeks Z(n) giraffe(f) pink(f)
'What Dan seeks is a pink giraffe.'

b. ma še-dan mexapes zot jirafa vruda **only de re**
what that-Dan seeks Z(f) giraffe(f) pink(f)
de re of 'What Dan seeks is a pink giraffe.'

To represent the two reading, I follow Sharvit's (1999) analysis of opacity connectivity in English which is based on Zimmermann's (1993) theory of intensional predicates: the opaque reading is obtained as an equation of properties, as in (28a), and the *de re* reading by scoping out the indefinite, as in (28b), yielding an equation between the entity denoted by the FR and an entity variable (for simplicity, I use the iota operator in FRs, not the Max(imality) operator – see Sharvit for discussion).

(28) a. de dicto (type <s,<e,t>>) $\iota\pi(\text{seek'}(d,\pi)) = {}^{\wedge}PG$
b. de re (type e) $\exists x(PG(x) \,\&\, \iota y(\text{seek'}(d,{}^{\wedge}\lambda z.z=y)) = x)$[7]

Both representations are available for the cross-categorial neutral pronZ accounting for the ambiguity of (28a), i.e. for why it exhibits opacity connectivity. Agreeing pronZ which is restricted to equating entities, only has the meaning in (29b), explaining why it is not "connected" with respect to opacity. Further support for this analysis comes from the behavior of the post-copular idiom in (29) which only allows for neutral pronZ.

[7] Note that any way of assigning the post-copular indefinite an individual denotation will explain why we only get the *de re* reading with agreeing pronZ. One alternative to scoping out the indefinite is using a choice function (Reinhart 1997) which would yield: $\exists g(CF(g) \,\&\, \iota y[\text{seek'}(d,{}^{\wedge}\lambda z.z=y)] = g(PG))$.

(29) ma še-dan mexapes ze /#zot maxat be-aremat šaxat
what that-Dan seeks Z(n)/Z(f) needle in-stack hay
'What Dan seeks is a needle in a haystack.'

Since the idiomatic expression *maxat be-aremat šaxat* 'a needle in a haystack' does not have a denotation at type e, it cannot be equated by agreeing pronZ. If we force the agreeing pronZ version, we lose the idiomatic reading and get a funny reading where Dan is looking for an actual needle in a haystack.

Crucially, this analysis does not involve any special mechanism for deriving opacity connectivity: the "connected" *de dicto* reading is a result of assuming Zimmermann's (1993) theory of intensional predicates, which is required for representing opacity in other environments, together with the cross-categorial nature of neutral pronZ. The absence of this reading from the agreeing pronZ pseudocleft follows directly from its type restriction. If, however, we wish to derive connectivity by positing c-command relation between the intensional verb and its object at some level of representation using some Operation C, we have to assume that it applies in neutral pronZ pseudoclefts but not in agreeing pronZ pseudoclefts. While this will yield the right result for opacity connectivity, it will make the wrong prediction for other connectivity effects, such as Bound Variable connectivity discussed in the next section.

3.2. *Bound variable connectivity*

We saw in section 1 that both types of pronZ pseudoclefts exhibit Bound Variable (BV) connectivity. This connectivity effect, originally noticed by Dahl (1981) and Hornstein (1984) for English, is the name for cases where a pronoun in the post-copular phrase is bound by a non-c-commanding quantified NP inside the FR. Consider (30), repeated from (6), which is possible with both types of pronZ.

(30) mi še-**[kol gever]$_i$** ohev ze /zot ima šel-**o$_i$**
who that-every man loves Z(n)/Z(f) mother of-his
'The person every man loves is his mother.'

Jacobson (1994) accounts for this connectivity effect using functions. The idea is that the phrases around the copula don't denote entities, but rather functions from entities into entities (type <e,e>), as in (31).

(31) $\iota f_{<e,e>}[\forall x(man'(x) \rightarrow love'(x, f(x)))] = \lambda y.\iota z[mother\text{-}of'(z,y)]$

The binding of the pronoun by the quantified NP is indirect, and it follows from the identity between the two functions: the (unique) function that maps every man to whoever he loves is identified with the function that maps every individual onto his mother, and this is equivalent to asserting that every man loves his mother.

The <e,e> functions are possible arguments for the cross-categorial neutral pronZ, accounting for the BV connectivity in the neutral pronZ version of (30). The fact that the agreeing pronZ version of (30) is also grammatical is prima facie unexpected, as we have seen that agreeing pronZ can only equate ("real world") entities and not arguments of higher type (see again example 26). I propose that these functions can be an argument of agreeing pronZ because their range is "real world" entities. More generally, this means that shifting the denotation of the noun into a functional denotation of type <e,X> is nothing more than a formal extension of its denotation at type X. Accordingly, agreeing pronZ can equate functions into "real world" entities in addition to simple "real world" entities.

Evidence for this proposal comes from other cases of BV connectivity where the post-copular NP does not denote a "real world" entity. (32) shows that when the post-copular NP is an event-denoting entity as in (32) only neutral pronZ is possible. The same

is true when the post-copular NP is a proposition-like entity as in (33).

(32) ma še-**[kol marce]$_i$** hicliax limno'a ze /*zot ha'ataka ba-mivxan šel-**o$_i$**
what that-every lecturer managed to-prevent Z(n)/Z(f) copying in-the-test of-his
'What every lecturer managed to prevent was copying on his exam.'

(33) ma še-**[kol student]$_i$** daxa ze /*zot te'ana bidvar heyot-**o$_i$** aclan
what that-every student rejected Z(n)/Z(f) claim(f) about being-his lazy
'What every student rejected was a claim about his being lazy.'

The BV reading in the neutral pronZ pseudoclefts is represented using an equation of functions from individuals into event-denoting entities in (32) or into proposition-like entities in (33). Such functions are not possible arguments for agreeing pronZ, because the basic denotation of these nouns cannot be equated by agreeing pronZ. The contrast between (30) and (32-33) demonstrates that agreeing pronZ cannot equate <u>all</u> <e,e> functions, but is restricted to those that are a formal extension of "real world" entities.

The functions used to derive BV connectivity are not particular to specificational pseudoclefts. Functional nouns and functional relative clauses are based on the analysis of functional questions and their answers in Groenendijk & Stokhof (1984), Engdahl (1986) and Chierchia (1993). (34) is an example of a functional question and its answer, both in English and in Hebrew.

(34) a. Q: Who does **[every man]$_i$** love? A: **His$_i$** mother.

b. Q: et mi **[kol gever]$_i$** ohev? A: et ima šel-**o$_i$**
Acc who every man loves Acc mother of-his
'Who does every man love?' 'His mother.'

These authors analyze the question as inquiring after the function of type <e,e> such that every man loves the output of that function for him. This is arrived at by assuming that the gap is of a complex nature – it contains a variable of type <e,e>. The answer to a

functional question is also functional – the nominal denotes the function into one's mother.

In Jacobson's variable-free semantics, the gap does not contains a variable. Instead, the functional denotation is achieved by shifting the meaning of the verb using a new type-shifting rule (the *z* rule). The functional nominal again contains no variables: the pronoun denotes the identity function on individuals, and hence the default denotation of such a nominal is functional. This apparatus, which is posited for dealing with functional questions and answers, suffices to derive BV connectivity in pseudoclefts[8] – the pre-copular FR denotes a maximal function due to the raised verb and the post-copular nominal is functional.

Sharvit (1999) adopts Jacobson's account of BV connectivity, but her derivation is closer to Groenendijk & Stokhof, Engdahl and Chierchia in that she uses variables. In particular, in her analysis the pre-copular FR denotes a function of type <e,e> because the verb is raised to take a variable of this type as its complement. The post-copular functional noun requires a type shifting rule that shifts the noun into a functional denotation. Note that this rule is required independent of specificational sentences: it is used in order to account for functional answers in any system that uses variables.

Whichever analysis we adopt to account for BV connectivity, the functional denotations are independently motivated by functional questions and their answers. The

---

[8] Jacobson needs an additional rule in order to deal with non-pseudoclefted specificational sentences – one which turns the denotation of a lexical noun into a functional one. This is in order for it to have the same type as the set of functions denoted by the relative clause with which it will be intersected. Note that Sharvit, who uses variables, postulates this rule for all functional nouns.

only element that is found in copular sentences and not in questions is the verb *be*, which we already assume to be cross-categorial[9]. Importantly, the same apparatus is also used for relative clauses with quantifiers (Jacobson 1994, 1999, Sharvit 1997) and BV connectivity in predicational sentences (Sharvit 1997, 1999).

How can we account for BV connectivity using Operation C? First, we would have to assume that Operation C applies both in neutral pronZ and in agreeing pronZ pseudoclefts. This crucially differs from what we had to assume in order to derive the right pattern for opacity connectivity, namely, that Operation C applies only in neutral pronZ pseudoclefts. It is unclear that these contradicting assumptions can be reconciled without ad-hoc stipulation.

Note, in addition, that deriving BV using Operation C does not mean that we can forgo functional nouns and functional relative clauses in the grammar. Thus, even if we can find a reason for why Operation C applies only in those agreeing pronZ pseudoclefts that exhibit BV connectivity, this option requires more apparatus than the analysis proposed here.

3.3. *Case connectivity and the accusative marker 'et'*

The third connectivity effect we saw in section 1 involved the Accusative marker *et*, which obligatorily marks all definite direct objects in Hebrew. (35), repeated from (3), shows that *et* must mark a post-copular definite NP when the gap inside the FR is in

---

[9] As noted by Sharvit, for non-pseudoclefted specificational sentences we also need to assume that *the* is cross-categorial. Again, this assumption is not specific to BV connectivity in specificational sentences.

object position, and *et* cannot occur on the post-copular definite NP when the gap inside the FR is in subject position.

(35) a. [ma še-kaninu ba-šuk _] ze *(et) ha-xulca ha-kxula
what that-we-bought in-the-market Z Acc the-shirt the-blue
'What we bought in the market is the blue shirt.'

b. [ma še _ nikra ba-tiyul] ze (*et) ha-xulca ha-kxula
what that tore in-the-trip Z Acc the-shirt the-blue
'What tore during the trip was the blue shirt.'

Unlike the connectivity effects we saw up to this point, the Case status of a phrase is not encoded in its semantic type. So in order to account for Case connectivity, we have to introduce syntactic features into the representation (this connectivity effect has not been previously discussed within the direct compositional approach). The idea is that in the same way that the two phrases around the copula must be of the same semantic type in order to be equated, they must also bear the same syntactic features.

Specifically, the post-copular NP in the grammatical version of (35a) bears Accusative Case features by virtue of being a definite NP marked by *et*. The post-copular NP in the grammatical version of (35b), on the other hand, is unambiguously Nominative, as it is a definite NP <u>not</u> marked by *et*. Note, importantly, that some NPs may be underspecified for Case if their form stays the same in different syntactic positions (e.g. indefinite NPs in Hebrew do not bear Case marking in either subject or object positions).

The FR bears the Case features of the *wh*-word by which it is introduced, which are identical to the Case features of the gap: Accusative in (35a) and Nominative in (35b). The actual syntactic mechanism for marking the FR for Case features is not crucial here, but note that the idea that FRs as a whole bear Case is not new. It is known from the fact that in many languages FRs exhibit "matching effects": the Case of the FR as a whole (= the Case of its gap) has to match the Case assigned to the argument position in which the

FR occurs – see e.g. Groos & van Riemsdijk (1981) and Grosu (1994) for examples and analysis. Getting back to Case connectivity, a pronZ pseudocleft will only be grammatical if both phrases around the copula bear the same Case features: Accusative in (35a) and Nominative in (35b).

The connectivity effect of *et* marking shows up only in neutral pronZ pseudoclefts. (36), repeated from (5), illustrates the contrast in *et* marking between neutral and agreeing pronZ pseudoclefts (see also 24b-c).

(36) a. ma še-kaninu ba-šuk ze *(et) ha-xulca ha-kxula
what that-we-bought in-the-market Z(n) Acc the-shirt(f) the-blue(f)
b. ma še-kaninu ba-šuk zot (*et) ha-xulca ha-kxula
what that-we-bought in-the-market Z(f) Acc the-shirt(f) the-blue(f)
both: 'What we bought in the market was the blue shirt.'

This contrast seems to suggest that unlike neutral pronZ pseudoclefts, agreeing pronZ pseudoclefts do not exhibit Case connectivity. However, when the post-copular NP is (unambiguously) Nominative, we do find Case connectivity with agreeing pronZ. In (37), the post-copular Nominative pronoun is only possible when the gap in the FR is in subject position.

(37) a. mi še-hirgiza oti zot hi
who that-annoyed me Z(f) she(Nom)
'The person that annoyed me was her.'

b. *mi še-dan ohev zot hi
who that-Dan loves Z(f) she(Nom)
'The person Dan loves is her.'

We must therefore conclude that agreeing pronZ pseudoclefts do exhibit Case connectivity. But this leaves unanswered the question of why agreeing pronZ pseudoclefts do not allow for *et* marking.

An interesting analysis of *et* is presented by Danon (2001), who argues that *et* is a type-shifting operator that lifts its argument from a type e denotation into the type of

generalized quantifiers (type <<e,t>,t>). This analysis of *et* can explain why *et* cannot occur with arguments of agreeing pronZ: an *et* marked phrase is a generalized quantifier, yet agreeing pronZ can only equate entities. Some support for relating the occurrence of *et* to the type of the arguments comes from pseudoclefts where the pre-copular FR is introduced by *mi* 'who'. Independent of the type of pronZ, *et* cannot mark the post-copular definite NP, even though the gap in the FR is in object position.

(38) mi še-dan ohev ze /zot (*et) ha-šxena mimul
who that-Dan loves Z(n)/Z(f) Acc the-neighbor(f) from-across
'The person Dan loves is the next-door-neighbor.'

The fact that *et* marking of the post-copular NP depends on the *wh*-word that introduces the FR and not just on the position of the gap in the FR supports the idea that *et* is more than just a Case marker. In particular, if an *et* marked phrase denotes a generalized quantifier, the fact that it cannot be equated with a *mi* 'who' FR indicates that such a FR cannot denote a generalized quantifier. It seems prima facie plausible that the fact that *mi* 'who' refers to human individuals may restrict the type of denotation for its FRs. Note in conclusion that analyzing *et* as a type-shifting operator does not replace the syntactic feature analysis for Case connectivity; the type analysis alone would not be able to account for cases where the post-copular NP is indefinite, or for cases like (37) where the Case of the arguments is Nominative.

To conclude the discussion of *et* marking and Case connectivity, let us compare the analysis presented here with deriving these connectivity effects using Operation C. First, to derive Case connectivity, we would have to assume that Operation C applies in both types of pronZ pseudoclefts. One problematic consequence of this assumption would be that it would not predict the different behavior of Case connectivity and *et* marking, since

the two always co-occur in simple sentences. The analysis presented here can naturally account for such a mixed pattern because different parts of the representation "conspire" to yield connectivity: syntactic features for Case and type for *et* marking.

From a broader perspective, applying Operation C in both types of pronZ pseudoclefts will not predict the right pattern for other connectivity effects that show up just in neutral pronZ pseudoclefts, such as opacity (and some cases of BV connectivity). It is unclear how the different patterns found for the different connectivity effects can all be accounted for by one operation.

3.4. *Agreement connectivity*

A second connectivity effect that is not found in English pertains to subject-verb agreement across the copula. (39) illustrates that a predicate in Hebrew must agree with its subject in gender (and number): the feminine subject in (39a) requires a feminine verb and the masculine subject in (39b) requires a masculine verb.

(39) a. ha-šxena *loke'ax / lokaxat et rut me-ha-gan
the-neighbor(f) takes(m) / takes(f) Acc Ruth from-the-kindergarten
'The neighbor takes Ruth from the kindergarten.'

b. ha-šaxen loke'ax / *lokaxat et rut me-ha-gan
the-neighbor(m) takes(m) / takes(f) Acc Ruth from-the-kindergarten
'The neighbor takes Ruth from the kindergarten.'

It is not the case, however, that the grammatical gender of the NP subject determines the gender on the verb. In (40), for instance, the grammatical gender of the NP subject is masculine, yet the verb can be either masculine or feminine, depending on the gender of the referent.

(40) roš ha-xug axrai /axra'it al kabalat studentim xadašim
head(m) the-department in charge(m)/ (f) on admitting students new
'The department chair is in charge of admitting new students.'

These data suggest that agreement in Hebrew is semantic, as the verb agrees with the gender of the individuals denoted by the NPs and not with the gender of the NPs themselves. To account for agreement in Hebrew, I adopt Dowty & Jacobson's (1989) semantic theory of agreement. The model they propose marks gender by sorting the domain of entities for natural and non-natural gender. In such a model, entities are marked for gender, and predicates (which are sets of entities – type <e,t>) are marked for gender indirectly/ The set of entities denoted by *loke'ax* 'takes (m)' contains only masculine entities, and it is distinct from the set denoted by *lokaxat* 'takes (f)' which contains just feminine entities.

This section discusses two kinds of subject-verb agreement across the copula: one where the "predicate" in the post-copular position agrees with the subject inside the FR, and one where the verb inside the FR agrees with the "subject" in the post-copular position. Agreement connectivity has not been accounted for within the direct compositional approach to connectivity which has been developed only for English. For Hebrew, I will present a semantic account of agreement connectivity. But in languages where agreement is syntactic[10] it is possible to have a syntactic analysis of connectivity, along the lines of the analysis given for Case connectivity in section 3.3.

---

[10] One such language is Italian, where a masculine NP always triggers masculine agreement on the verb, even when it is used to refer to a feminine individual (thanks to Ivano Caponigro for pointing this out).

(i) ITALIAN
Il presidente è seriamente preoccupato /*preoccupata
the president is seriously worried(m) /*worried(f)
'The president is very worried.'

3.4.1. A post-copular "predicate"

The first kind of agreement connectivity is observed in pseudoclefts where the gap inside the FR is in predicate position and the post-copular phrase is a predicate. In these pseudoclefts, the post-copular predicate agrees in gender (and number) with the subject inside the FR. In (41a), where the subject inside the FR is feminine, the post-copular adjectival predicate must also be feminine, and in (41b), where the subject inside the FR is masculine, the post-copular adjectival predicate must also be masculine. Note that only neutral pronZ is possible here as this is an equation of predicates, not entities.

(41) a. ma še-**rut** hayta ze *mo'il la-xevra **/ mo'ila la-xevra**
what that-Ruth was(f) Z(n) *helpful(m) to-the-society / helpful(f) to-the-society
'What Ruth was was helpful to society.'

b. ma še-**dan** haya ze **mo'il la-xevra** /* mo'ila la-xevra
what that-Dan was(m) Z(n) helpful(m) to-the-society/* helpful(f) to-the-society
'What Dan was was helpful to society.'

In order to account for the connectivity effect, we need to assume that both phrases around the copula bear the same gender. That is, the predicate and the FR in (41a) should somehow both be marked as feminine and the predicate and FR in (41b) as masculine. Note, however, that when these FRs occur in subject position of a predicational sentence, they both trigger masculine agreement on the verb – this shows that the same FRs can have two different denotations.

(42) a. ma še-rut hayta hirgiz /*hirgiza oti
what that-Ruth was(f) annoyed(m)/annoyed(f) me
'What Ruth was annoyed me.'

b. ma še-dan haya hirgiz /*hirgiza oti
what that-Dan was(m) annoyed(m)/annoyed(f) me
'What Dan was annoyed me.'

I will therefore assume that when occurring in a neutral pronZ pseudocleft in (41) the FRs denote predicates (type <e,t>), whereas in the (referential) subject position of (42) they denote the intensional correlate of this predicate, i.e. the denotation is of type <s, <e,t>>. The difference in gender between the two denotations is straightforward in the model described above. Being a set of gender bearing entities, a predicate is masculine if it is a set of masculine individuals and it is feminine if it is a set of feminine individuals. That is, while predicates are not directly marked for gender, they bear gender by virtue of being sets of entities that bear gender. But the intensional correlates are not marked for gender: the gender of individuals is not necessarily constant across possible worlds[11]. The FRs in (42) are thus genderless and as such they trigger default (=masculine) gender on the verb.

We can now account for agreement connectivity. In (41), the post-copular predicate *mo'il la-xevra* 'helpful(m) to society' denotes a set of masculine entities and the post-copular predicate *mo'ila la-xevra* 'helpful(f) to society' denotes a set of feminine entities. The pre-copular FRs in (41) are also predicates: the (contextually unique) predicate that Ruth has in (41a) and the (contextually unique) predicate that Dan has in (41b). Since Ruth denotes a feminine entity and Dan denotes a masculine entity, the predicates consist of masculine and feminine entities respectively. The connectivity effect is a result of the equation semantics: in order for the two sets to be identical, they must both contain entities of the same gender.

---

[11] Another possibility would be to assume that the FRs in (42) refer to property-like individuals – a subtype of type e and assume that this sort, which is distinct from "real world" individuals, is not sorted for gender. We will see in section 3.5 that this is not supported by subject-predicate agreement data.

3.4.2. A post-copular "subject"

The second kind of agreement connectivity is found when the gap in the FR is in subject position and the post-copular phrase is an NP. This connectivity effect yields a pattern that we have not seen with the other connectivity effects: it is present in agreeing pronZ pseudoclefts and absent from neutral pronZ pseudoclefts (which show "anti connectivity" with respect to such agreement). (43) illustrates agreement connectivity in agreeing pronZ pseudoclefts: the verb inside the FR must agree with the post-copular feminine NP. (44) shows that neutral pronZ pseudoclefts exhibit the opposite pattern: the verb inside the FR must be masculine, even though the post-copular NP is feminine – this is a case of anti-connectivity[12].

(43) a. mi še-**lokaxat** et rut me-ha-gan **zot** ha-šxena
who that-takes(f) Acc Ruth from-the-kindergarten Z(f) the-neighbor(f)
b. *mi še-loke'ax et rut me-ha-gan zot ha-šxena
who that-takes(m) Acc Ruth from-the-kindergarten Z(f) the-neighbor(f)
both: 'The person that takes Ruth from the kindergarten is the neighbor.'

(44) a. *mi še-lokaxat et rut me-ha-gan ze ha-šxena
who that-takes(f) Acc Ruth from-the-kindergarten Z(n) the-neighbor(f)
b. mi še-**loke'ax** et rut me-ha-gan **ze** ha-šxena
who that-takes(m) Acc Ruth from-the-kindergarten Z(n) the-neighbor(f)
both: 'The person that takes Ruth from the kindergarten is the neighbor.'

The agreement pattern of agreeing pronZ pseudoclefts follows from the fact that this copula only equates "real world" entities. Since the post-copular NP denotes a feminine entity, the entity denoted by the FR must also be feminine. This is supported by subject-verb agreement in (45) – we see that when the FR refers to a person, the gender of the

[12] The FRs used here are introduced by *mi* 'who', because when a FR is introduced by *ma* 'what' the verb in the FR is always masculine-singular, even if the actual referent is feminine, e.g. if the FR in (i) refers to a banana, which is feminine in Hebrew.
(i) ma še-kilkel /*kilkela li et ha-te'avon
what that-ruined(m)/*ruined(f) to-me Acc the-appetite 'What ruined my appetite'

verb inside the *mi* 'who' FR determines the gender of that person.

(45) a. mi še-**lokaxat** et rut me-ha-gan *hu nexmad **/ hi nexmada**
who that-takes(f) Acc Ruth from-the-kindergarten is(m) nice(m) / is(f) nice(f)
b. mi še-**loke'ax** et rut me-ha-gan **hu nexmad** /*hi nexmada
who that-takes(m) Acc Ruth from-the-kindergarten is(m) nice(m) /is(f) nice(f)
both: 'The person that takes Ruth from the kindergarten is nice.'

As a result, only a FR with a feminine verb is possible in (43).

As for neutral pronZ, the ungrammaticality of (44a) indicates that a neutral pronZ pseudocleft cannot be an equation of "real world" entities (the phrases around the copula could in principle denote feminine individuals as in 43a). This forces us to adapt the cross-categorial characterization we have been assuming for neutral pronZ: (44a) reveals that neutral pronZ is not completely cross-categorial, but is restricted to non-"real world"-entities – I discuss the implications of this revised generalization in section 3.5. In the grammatical (44b), the fact that the two phrases around the copula do not seem to bear the same gender suggests that this is an equation of semantic objects that are not marked for gender. Two options are properties (type <s,<e,t>>) and individual concepts (type <s,e>): being intensional, these are not marked for gender (see discussion of 42 above). The reason the verb inside the FR is marked masculine is that this is the default gender marking.

This pattern of agreement connectivity follows from the type restrictions that agreeing and neutral pronZ pose on their arguments. The restriction of agreeing pronZ to "real world" entities forces both arguments to denote entities of the same gender, i.e. to exhibit "connectivity" – it is the same restriction that blocks other connectivity effects. Our revised generalization, stating that neutral pronZ cannot equate "real world" entities blocks the "connected" reading here, but does not block other connectivity effects.

We see once again that this mixed pattern cannot be arrived at using Operation C. Applying Operation C in neutral pronZ pseudoclefts will yield the wrong pattern for "subject" agreement connectivity, and applying it in agreeing pronZ pseudoclefts will derive the wrong pattern for opacity, (some cases of) Bound Variable connectivity, and *et* marking.

3.5. *Interlude: agreement of pronZ*

It was argued in the beginning of section 3 that agreeing pronZ can only equate "real world" entities, and we saw how this restriction on the type of the arguments accounts for the connectivity pattern observed in agreeing pronZ pseudoclefts. For neutral pronZ, we have assumed that it is cross-categorial like English *be*, but in order to account for the anti-connectivity effect found with respect to "subject" agreement connectivity (section 3.4.2) we had to revise our generalization and restrict its arguments to non-"real world" entities. As a result, those neutral pronZ pseudoclefts that may at first seem to be an equation of "real world" entities, such as (27a) or (38), must be reanalyzed at a higher type. Such reanalysis is possible in any system that includes general type-shifting rules.

In the model presented here, where entities are marked for gender (see again section 3.4), it would be attractive to reanalyze the two types of pronZ as a single pronZ that reflects the gender of its arguments. The idea is that objects in the model that bear gender, i.e. entities, would reflect their gender on pronZ, resulting in a copula which may be masculine or feminine. By contrast, objects that do not bear gender, i.e. higher order objects and other non-entities, would reflect their gender on the copula, but since there is no gender, the copula will be marked with default gender – masculine. That is, instead of having neutral pronZ be identical to the masculine agreeing pronZ, we collapse them into

one element that reflects either real masculine or default gender.

This alternative analysis of pronZ is attractive for two reasons. First, it reduces the number of elements in the grammar. But, more importantly, it gives an explanation for the agreement behavior of pronZ. Unfortunately, it faces a problem in the treatment of type e entities that are not "real world" entities. We saw in the beginning of section 3 that NPs denoting such entities (property-like, proposition-like and event-denoting entities) are only possible with neutral pronZ (see again the examples in 27), so under the unified characterization of pronZ such entities are predicted not to bear any gender features.

This prediction, however, contradicts what we observe in other environments. In particular, when these NPs occur in subject position of a simple sentence, the verb shows agreement in accordance with their grammatical gender.

(46) a. te'ana bidvar šxitut ha-ma'arexet *ho'ala / ho'alta ba-yešiva
claim(f) about corruption the-system was-raised(m)/was-raised(f) in-the-meeting
'A claim about corruption in the system was raised in the meeting.'

b. acirat gšamim ba-xoref ha-karov *yesaken/ tesaken et ha-yevul
stopping(f) rains in-the-winter the-close will-risk(m)/will-risk(f) Acc the-crop
'A stopping of the rain next winter will risk the crop.'

At this point it is unclear if and how the two pieces of data can be reconciled. Therefore, I reject the unified characterization of pronZ and maintain the hypothesis that there are two types of pronZ. It is my hope that further research on other gender agreement patterns in Hebrew will shed light on the appropriate gender marking in the semantic model and will allow for a unified analysis of pronZ.

3.6. *Binding theory connectivity*

The standard Binding Theory (Chomsky 1981, 1986) is an account of the distribution of anaphoric elements (reflexives, pronouns, and proper names) which is stated in term of

structural relations between the antecedent and the anaphor within a minimal syntactic domain. In specificational sentences, however, these anaphoric elements occur without the structural configuration – these are cases of Binding Theory connectivity. Accounting for these cases within the null hypothesis calls for a theory of the distribution of anaphors that does not rely on syntactic configuration. Since developing such a theory is beyond the scope of the dissertation, this section concentrates on presenting the Binding Theory connectivity patterns in agreeing and neutral pronZ pseudoclefts, along with some preliminary suggestions for an analysis adopted from Jacobson (1994) and Sharvit (1999).

3.6.1. Principle A connectivity

The first connectivity effect we saw in this chapter was the binding of a reflexive across the copula. In this example, repeated here as (47), the reflexive is embedded in a post-copular predicate (so only neutral pronZ is possible).

(47) ma še-**dan** haya ze mesukan le-**acmo**
what that-Dan was Z(n) dangerous to-himself
'What Dan was was dangerous to himself.'

An account of this connectivity effect that does not rely on structural relations between the reflexive and its antecedent was proposed in Jacobson (1994). Jacobson adopts the approach of Bach & Partee (1980) and Szabolcsi (1987) (see also Reinhart & Reuland 1993) who take predicates with a reflexive complement to denote reflexive predicates. Authors differ in how they arrive at the reflexive predicate, but they all assume that in this environment the reflexive is an argument reducer. In Jacobson's variable-free semantics, for instance, the reflexive denotes the identity function on entities which forces the verb to raise into a relation between <e,e> functions and entities. Adopting this

analysis for the Hebrew reflexive *acmo* yields the following meaning for (47): the (unique) property that Dan has is the property of being dangerous to oneself, which is equivalent to Dan's being dangerous to himself.

(48) $\iota P[P(d)] = \lambda x.\text{dangerous-for}'(x,x)$

There are, however, cases of Principle A connectivity where the reflexive is not embedded within a predicate. Consider (49): the neutral pronZ version in (49a) allows for binding of a reflexive (which must also be *et* marked), and agreeing pronZ in (49b) blocks such indirect binding (independent of *et* marking).

(49) a. ma še-**rut** haxi ohevet ze *(et) **acma**
what that-Ruth most loves Z(n) Acc herself
'What Ruth loves most is herself.'

b. *ma še-**rut** haxi ohevet zot (et) **acma**
what that-Ruth most loves Z(f) Acc herself

Sharvit (1999), who adopts Jacobson's analysis of Principle A connectivity for examples like (47), points out that if we apply Jacobson's analysis of the reflexive as the identity function on entities to (49a), the sentence will have the bound reading. In particular, the FR denotes the (unique) function that maps Ruth to whoever she loves (parallel to functional FRs in Bound Variable connectivity), and the post-copular reflexive denotes the identity function on entities. The equation of these functions is logically equivalent to asserting that Ruth loves herself.

(50) $\iota f[\text{love}'(r,f(r))] = \lambda x.x$

Note that if the Accusative marker *et* is indeed a type-lifting operator, the meaning assigned to the reflexive, and subsequently the meaning of the FR, would have to be lifted to the type of generalized quantifiers.

Pollard & Sag (1983, 1992) propose a reformulation of Principle A in which a reflexive is bound only if it is an argument of a predicate with a higher argument slot. That is, if a reflexive occurs in a position which does not have a higher argument slot, it will be "exempt" and its referent will be determined by processing and discourse factors. In order for the reflexive in (49a) to has this status, we have to argue that it is not an argument of a predicate that has an argument slot. The relevant predicate here is pronZ. Assuming it is a 'be of identity', it does actually have a higher argument slot, i.e. the pre-copular position. Thus, if we wish to adopt Pollard & Sag's reformulation, we would have to argue that this predicate is nonetheless different from other predicates and such that the reflexive is exempt. While this of course requires independent evidence, it seems plausible for pronZ which, as a pronominal copula, is very different from other predicate in Hebrew. Note that this is less straightforward for the English *be*. It is interesting to mention in this context that Pollard & Sag (1992) themselves mention the fact that their Principle A does not render the post-copular position of a specificational sentence "exempt" as a weakness of the analysis – see Pollard & Sag (1992, fn. 40). Settling this matter will be left for future research.

In conclusion, it should be pointed out that the ungrammaticality of the agreeing pronZ (49b) is independent of Principle A connectivity. To illustrate this, consider reflexive connectivity with *mi* 'who' FRs which block *et* marking on the post-copular NP (see again example 38).

(51) *mi še-rut haxi ohevet ze /zot acma
who that-Ruth most loves Z(n)/Z(f) herself
'The person Ruth loves most is herself.'

Assuming that the post-copular reflexive denotes an <e,e> function, the ungrammaticality of both versions of (51) cannot be due to the denotation of a *mi* FR, which can in principle have a denotation at this type (see again section 3.2). In addition, the ungrammaticality of the neutral pronZ version cannot be attributed to the absence of reflexive connectivity, which has been shown to be possible in (49a). Therefore, one has to conclude that the source for the ungrammaticality of (51) is the ban on *et* marking. What this example shows is that the Hebrew reflexive *acmo*/*acma* cannot occur without *et*. This brings us back to the ungrammaticality of (49b): this is because an agreeing pronZ pseudocleft does not allow for *et* marking on its post-copular NP. This fact prevents us from determining whether agreeing pronZ pseudoclefts exhibit reflexive connectivity.

3.6.2. Principle B connectivity

The second connectivity effect covered by Binding Theory is binding of a pronoun. Both the neutral and the agreeing pronZ pseudocleft in (52) are ungrammatical.

(52) a. *ma še-**rut$_i$** haxi ohevet ze **ot-a$_i$**
what that-Ruth most loves Z(n) Acc-her
b. *ma še-**rut$_i$** haxi ohevet zot **ot-a$_i$**
what that-Ruth most loves Z(f) Acc-her
both: '*What Ruth$_i$ loves most is her$_i$.'

Accusative pronouns in Hebrew are pronominal suffixes on *et*, so the ungrammaticality of (52b) is due to the ban on *et* in agreeing pronZ pseudoclefts; therefore, the ungrammaticality may be unrelated to Principle B connectivity. If we were to replace the Accusative pronoun with a Nominative pronoun (which is not a suffix on *et*), the pseudocleft would be ungrammatical due to Case mismatch. Therefore, there is no way to

determine whether agreeing pronZ pseudoclefts exhibit Principle B connectivity (note that *mi* FRs will raise the same problems).

Accounting for the connectivity effect in (52a) within the direct compositional approach requires an alternative to the standard Binding Theory. Such a theory is currently unavailable, but Sharvit (1999) presents a interesting line of thought, adopting an idea from Grodzinsky & Reinhart (1993). Very briefly, Grodzinsky & Reinhart suggest that a coreferential reading is blocked if there is an alternative grammatically marked bound reading ("Rule-I"). If we apply this idea to (52a), its ungrammaticality would be accounted for by the existence of the bound reading with a reflexive in (49a). I find this to be a promising line, but a complete theory of the distribution of pronouns is beyond the scope of this dissertation.

3.6.3. Principle C connectivity

Our last connectivity effect deals with the distribution of proper names in the post-copular position. The data in (53-54) shows that both the neutral pronZ version and the agreeing pronZ version are ungrammatical, independent of whether the FR is introduced by *ma* 'what' or *mi* 'who'. Recall that the latter disallows *et* marking on the post-copular NP. In order to determine which of these are indeed instances of Principle C connectivity and which are ungrammatical for independent reasons, we replace the subject pronoun inside the FR with a proper name, e.g. *rina* (female).

(53) a. *ma še-**hi$_i$** haxi ohevet ze et **rut$_i$** (OK with *rina*)
what that-she most loves Z(n) Acc Ruth
b. *ma še-**hi$_i$** haxi ohevet zot **rut$_i$** (* with *rina*)
what that-she most loves Z(f) Ruth
both: '*What she$_i$ loves most is Ruth$_i$.'

(54) *mi še-**hi**$_{i}$ haxi ohevet ze / zot **rut**$_{i}$ (OK with *rina*)
who that-she most loves Z(n)/Z(f) Ruth
'*The person she$_{i}$ loves most is Ruth$_{i}$.'

The fact that (53b) is ungrammatical with a proper name shows that its ungrammaticality is not a result of Principle C connectivity. But (53a) and (54) are instances of Principle C connectivity, i.e. this connectivity effect is found in both types of pronZ pseudoclefts. The fact that agreeing pronZ pseudoclefts exhibit Principle C connectivity is important, as we could reach no conclusion with respect to Principle A and Principle B connectivity. Since accounting for Principle C connectivity within the direct compositional approach requires a non-structural theory of the distribution of proper names, this issue is left for future research.

3.7. *Summary*

We saw in this section that neutral pronZ pseudoclefts exhibit most connectivity effects, whereas agreeing pronZ pseudoclefts block many connectivity effects. The pattern is summarized in Table 3.1.

Connectivity Effect	Neutral pronZ	Agreeing pronZ
Opacity	✓ (27a)	× (27b)
Bound Variable:		
"real world" entities	✓ (30)	✓ (30)
property-like entities	✓ (32)	× (32)
proposition-like entities	✓ (33)	× (33)
Case	✓ (35)	✓ (37)
et marking	✓ (36a) *ma* × (38) *mi*	× (36b, 38)
Agreement: "predicate"	✓ (41)	× (25c,26b)
"subject"	× (44)	✓ (43)
Binding Theory: Principle A	✓ (49a)	? (49b)
Principle B	✓ (52a)	? (52b)
Principle C	✓ (53a,54)	✓ (54)

Table 3.1: connectivity effects with two pronZ copulas

We attributed the difference in their connectivity pattern to the different arguments they equate: agreeing pronZ is restricted to "real world" entities while neutral pronZ can equate non-"real world" entities. Although I have not provided a definition of "real world" entities, the fact that the grammar is sensitive to such a notion is by itself an important finding. I leave the exact definition of this notion for future work.

The connectivity effects were accounted for within the direct compositional approach (Higgins 1973), developed by Jacobson (1994) and Sharvit (1999). In particular, each

connectivity effect is reanalyzed such that its analysis does not rely on c-command; in specificational sentences the connectivity effects are a by-product of equation. The connectivity effects of opacity and Bound Variable were accounted for following the semantic theory of Jacobson (1994) and Sharvit (1999). Agreement connectivity was also accounted for semantically, by sorting the domain of entities for (natural and non-natural) gender. Accounting for Case connectivity required adding sameness of syntactic features to the requirement that (both types of) pronZ pose on the arguments. Adding syntactic features to the representation is a natural further step in developing the direct compositional approach, but a better understanding of the range of syntactic features that should be identical across the copula requires examining languages that have a wider range of 'syntactic' connectivity effects. The main problem with the current version of the direct compositional approach is dealing with Binding Theory connectivity – this requires a new theory of the distribution of anaphoric elements that does not rely on structural configuration. This issue calls for further research.

4. Alternative Analyses of Connectivity

Throughout the previous section I have pointed out that the Hebrew connectivity pattern constitutes evidence against deriving connectivity by a single grammatical operation special to specificational sentences. This section looks more closely at the predictions of specific analyses of connectivity.

4.1. *Positing c-command*

We first consider the family of analyses that account for connectivity effects by positing c-command between the material in the pre- and the post-copular position at an abstract level. Chapter 1 presented some of the different versions of this approach: copying the

post-copular material into the gap in the FR (Hornstein 1984), copying the FR to the post-copular position (Peters & Bach 1968), manipulating the pseudocleft into a simple sentence (Heycock & Kroch 1999). Chapter 2 presented two analyses within the 'Question plus Deletion' approach: den Dikken, Meinunger & Wilder (2000) and Schlenker (2003). Let us consider how these different theories can deal with the Hebrew pattern.

First, we look at neutral pronZ pseudoclefts. Accounting for connectivity effects in these sentences by positing c-command at an abstract level will give the wrong prediction for two connectivity effects: agreement connectivity with a post-copular "subject" and *et* marking connectivity with *mi* 'who' FRs. Since these connectivity effects do not form a natural class, it is unclear how one could block the derivation of just these two effects. Moreover, even if we assume ad-hoc that the c-command mechanism does not apply to sentences containing these two effects, this will yield the wrong prediction for pseudoclefts that exhibit more than one connectivity effect, e.g. the available Principle C connectivity and the blocked *et* marking connectivity in (54).

Next, we turn to the treatment of agreeing pronZ pseudoclefts. Applying some c-command mechanism in these sentences will make the wrong prediction for the effects that are not found; namely, opacity connectivity, (some cases of) Bound Variable connectivity and *et* marking connectivity. As with neutral pronZ pseudoclefts, applying the c-command mechanism to a subset of the sentences would be ad-hoc, and, more importantly, will not account for sentences which exhibit two connectivity effects that pattern differently, e.g. the available Bound Variable connectivity and the blocked *et* marking connectivity in (30). Alternatively, one could argue that the absence of many

connectivity effects means that the c-command mechanism should not apply in agreeing pronZ pseudoclefts; maybe they are not specificational (but see again section 2.2). However, if we adopt this analysis we will need a separate mechanism in order to account for the connectivity effects that are found in agreeing pronZ pseudoclefts.

In sum, the family of analyses that derives connectivity effects by relating them to the same effects in a simple sentence derives all connectivity effects at once – such an approach cannot deal with the Hebrew pattern where each type of pseudocleft exhibits a different range of connectivity effects.

4.2. *Question-answer pairs (without ellipsis)*

In the previous chapter we saw two version of the question-answer approach. First, there is the 'Question plus Deletion' approach, which derives connectivity effects in the post-copular obligatorily elided IP – we saw in the previous section why this approach cannot deal with the Hebrew connectivity pattern. In addition, we noted the possibility of a Question-Answer approach but without ellipsis, in which connectivity effects are accounted for by relating them to similar effects in question-answer pairs without positing a post-copular full answer. While at this point this approach is merely a logical possibility, it is worth examining whether such a theory could deal with the connectivity pattern in Hebrew.

Let us first consider how this theory would apply to neutral pronZ pseudoclefts. Interestingly, the anti-connectivity effects we saw with respect to gender marking on the verb inside a subject-gap FR (section 3.4.2), has a correlate in question-answer pairs. Compare (55a), repeated from (44b), with (55b).

(55) a. mi še-loke'ax / *lokaxat et rut me-ha-gan ze ha-šxena
who that-takes(m)/ takes(f) Acc Ruth from-the-kindergarten Z(n) the-neighbor(f)
'The person that takes Ruth from the kindergarten is the neighbor.'

b. Q: mi loke'ax et rut me-ha-gan? A: ha-šxena
who that-takes(m) Acc Ruth from-the-kindergarten the-neighbor(f)
'Who takes Ruth from the kindergarten?' 'The neighbor.'

This analogy accounts for the grammaticality of the masculine verb in (55), but in addition, we may use it to account for the ungrammaticality of the feminine verb in (55a): using a question with a feminine verb requires a special context and this may not be available for pseudoclefts. That is, the analogy to question-answer pairs accounts for agreement anti-connectivity.

The second connectivity effect that is not found in neutral pronZ pseudoclefts – *et* marking with *mi* 'who' FRs – does not have an exact correlate in question-answer pairs: while *et* is prohibited from the post-copular NP with *mi* 'who' FRs in (56a), it must mark the answer in (56b). Note that *et* must also mark the *mi* 'who' question, while the FR is n ever marked by *et*.

(56) a. mi še-dan ohev ze (*et) ha-šxena mimul
who that-Dan loves Z(n) Acc the-neighbor(f) from-across
'The person Dan loves is the next-door-neighbor.'

b. Q: *(et) mi dan ohev? A: *(et) ha-šxena mimul
Acc who Dan loves Acc the-neighbor from-across
'Who does Dan love?' 'The next-door-neighbor.'

Still, since *et* is a type shifting operator (see again section 3.3), there may be a technical reason for the difference between a pseudocleft and a real question-answer pair. That is, it might be possible to develop a question-answer analysis without ellipsis that will account for the connectivity pattern in neutral pronZ pseudoclefts.

However, it is unclear how this approach can account for the connectivity pattern in agreeing pronZ pseudoclefts. In particular, it is not clear how to block the missing

connectivity effects: opacity, (some cases of) Bound Variable and *et* marking. To avoid this problem, we can analyze agreeing pronZ pseudoclefts as something other than question-answer pairs. But then we would need an alternative account of the connectivity effects that are found in agreeing pronZ pseudoclefts: (some cases of) Bound Variable connectivity, Case connectivity, "subject" agreement connectivity and Principle C connectivity. In fact, this alternative account already exists. In particular, we could use the analysis presented in this chapter to account for the first three; a non-structural account of Principle C connectivity will have to be developed. But if the analysis presented here is needed to account for connectivity in agreeing pronZ pseudoclefts, it can be used "for free" to account for connectivity in neutral pronZ pseudoclefts. This would render the question-answer analysis superfluous.

In sum, while a Question-Answer approach without ellipsis is the most promising alternative to the direct compositional approach, at this point it is only a logical possibility. Moreover, even if such an analysis can be developed, it will be limited to neutral pronZ pseudoclefts – agreeing pronZ pseudoclefts still require specific analyses for the connectivity effects they exhibit.

5. CONCLUSIONS

This chapter presented a detailed direct compositional account of the complex connectivity pattern in Hebrew. First, this account demonstrates that the direct compositional approach can be applied to languages other than English (see also Ceccheto 2000, 2001 for a discussion of connectivity in Italian). Moreover, we see that the direct compositional approach, as opposed to other theories of connectivity, can handle a "mixed-bag" connectivity pattern such as the one found in Hebrew. This pattern

reveals that connectivity is not a uniform phenomenon and hence should not be analyzed using a single mechanism. Instead, connectivity is a cover term for a number of syntactic and semantic phenomena that tend to occur together, but may form different clusters depending on the circumstances.

The existence of a mixed pattern shows that there is not no one-to-one correlation between connectivity and specificational sentences (note that this conclusion stands independent of the label we assign to agreeing pronZ pseudoclefts). Therefore, connectivity cannot be used to determine whether a language has a specificational sentences. A language that may be relevant in this context is Modern Greek. Iatridou & Varlokosta (1998) show that no connectivity effects are found in Modern Greek, concluding that there are no specificational pseudoclefts in this language. In short, they follow Williams (1983) and Partee (1986) in that specificational pseudoclefts are an instance of inverse predication, and claim that FRs in Modern Greek cannot function as predicates because of the elements that introduce them. However, they identify a third type of pseudoclefts where the post-copular phrase is a CP ('CP-pseudoclefts') arguing that these are equatives rather than predicational or specificational. Under the analysis advocated here, it is possible that the reason there are no connectivity effects in Modern Greek is that the FRs do not have the "right" denotation to induce the different connectivity effects. In this picture, the equative 'CP-pseudoclefts' are simply specificational pseudoclefts that do not exhibit connectivity.

Taking together the Hebrew connectivity pattern presented in this chapter and the crosslinguistic data presented in the previous chapter we conclude that the direct compositional approach is the best account available for connectivity crosslinguistically.

Recall that this analysis assumes that specificational sentences are simply identity sentences. This by itself does not provide a definition of the specificational relation. In chapter 5, I propose that a certain pragmatic relation must hold between the pre- and the post-copular expressions in an identity sentence in order for it to be specificational. This relation concerns the relative 'discriminability' of the expressions. The next chapter is devoted to introducing this new notion of 'discriminability'.

Chapter 4
Introducing Discriminability

The previous chapters dealt with connectivity in specificational sentences. I have argued that connectivity effects are a by-product of the combinatorics of syntactic and semantic equation. This line of analysis, which assumes that specificational sentences are copular sentences that express identity does not by itself explain what makes these copular sentences special, i.e. what makes certain identity sentences specificational. Chapter 5 will address this issue; we will see that what characterizes specificational sentences is a certain pragmatic relation between the pre- and the post-copular expressions. This relation will concern the relative discriminability of the expressions. The goal of this chapter then is to introduce and discuss the notion of discriminability.

The data we use to introduce discriminability does not come from copular sentences. Rather, I present new data dealing with the choice of referring expressions. We will see that if both interlocutors in the discourse associate a certain entity with more than one potential referring expression, i.e. if there is information in the common ground that allows constructing more than one referring expression for the same entity, the choice between these referring expressions is not free. In particular, while the expressions may be equal in that they both leads to successful reference, there is a preference for an expression that best reflects the familiarity of the interlocutors with the entity. Greater familiarity with the entity means that the entity is better identified and can be better discriminated from other entities. Accordingly, an expression that reflects greater familiarity with the entity will be referred to as a more discriminate expression. Note that this approach depends on the idea that interlocutors are always in a state of partial

information, such that entities are not fully identified; during conversation, information may come in such that new entities are introduced and two entities that were assumed to be distinct may turn out to be one and the same.

The chapter is organized as follows. Section 1 presents the phenomenon: we will see that the choice among (definite) referring expressions depends on what information about the entity interlocutors share, i.e. on their combined knowledge states. Section 2 analyzes this pattern as a scalar phenomenon – we will also deal with two apparent counter-examples. I will argue that the choice among referring expressions is a device interlocutors use to try to keep their common grounds the same, i.e. to keep the context non-defective, in the sense of Stalnaker (1978). Section 3 presents Gundel, Hedberg & Zachariski's (1993) analysis of how the cognitive status of referents affects the choice of anaphoric expressions. We will see that my analysis shares some central properties with Gundel et al.'s analysis, even though they deal with two separate dimensions governing the choice of referring expressions. In the light of the two, I also discuss how other types of expressions – indefinites, pronouns and *-ever* FRs – fit into the picture.

1. Choosing a Referring Expression

This section presents a new pattern from the choice of referring expressions, considering four types of expressions – (i) proper names, (ii) definite descriptions in which the head is a lexically contentful noun, (iii) definite descriptions in which the head is "bleached" (such as *thing*, *place* or *person*), and (iv) standard FRs, i.e. FRs introduced by simple *wh*-words (*–ever* FRs will be discussed in section 3.4.). We will see that when more than one expression is available to refer to a certain entity, the different expressions are not of equal status – there is a preference for using an expression that encodes certain types of

information about the entity. This pattern will only be relevant (i) when the potential expressions are based on mutual knowledge, i.e. when the information that would allow constructing these expressions is available in the common ground, and (ii) when the expressions are used referentially, in the sense of Donnellan (1966). The two factors will be illustrated in the next section.

1.1. *Referring to individuals in English (the lasagna example I)*

Our first example considers the relative status of proper names and definite descriptions with a contentful noun – *neighbor*. We begin by considering a context of a potluck party. Let us assume that one of the guests is the next-door neighbor Giacomo who brought lasagna. The hostess and a guest are eating some of the delicious lasagna, and the guest inquires.

(1) Q Who brought the lasagna?

A1 Giacomo.
A2 The next-door neighbor.

The context of interest is such where the common ground contains information that would allow to refer to the individual in question using either the proper name in A1 or the description in A2. In this context, the information that Giacomo is the next-door neighbor is mutual, in the sense of Clark & Marshall (1981): each of the interlocutors knows this piece of information and they also know that it is mutually known. While in this context both descriptions can lead to successful reference, they do not have the same status for answering (1): the proper name is appropriate, whereas the definite description is odd as an answer – to feel the effect of oddness, I encourage the reader to change the name used here to their own neighbor's name. This asymmetry is unexpected if the only

goal of using the expression was to refer successfully, as both expressions can pick out the entity successfully.

It is important to point out that the asymmetry is found only when the information is mutual. If the common ground does not contain information that would allow constructing the two expressions, this pattern will not be found. For example, if the speaker doesn't know the name of the next-door neighbor, she will have no choice but to use the description. If, alternatively, she does not know that the two expressions are equi-extensional, the speaker will use whichever expression is associated with the entity she wishes to pick out. In addition, if the speaker is assuming that the addressee doesn't know that Giacomo is the next-door neighbor, she will use whatever expression can be constructed based on the information in the common ground – this may be the description. Note that, for ease of reference, I refer to a feminine speaker and a masculine addressee throughout this chapter.

While (1) shows that the name is clearly favored in contexts where both names and descriptions are available, I wish to draw attention to a putative counter-example. Consider a situation where the speaker does not have a certain entity in mind that she wants to talk about, but rather wishes to pick out an entity by virtue of its having a certain property. In particular, she wishes to pick out her neighbor, regardless of who this person is. Such use is not referential but rather attributive, in the sense of Donnellan (1966). (2) is an example of a question that will naturally be answered attributively because it inquires about a role.

(2) Q Who has to take out the garbage this week?

A1 Giacomo.
A2 The next-door neighbor.

Suppose the tenants of the apartments in a house took turns in taking the trash cans out to the sidewalk on the day of the weekly waste collection. In this case, the proper name and the definite description are not in competition: the description can be used attributively and the proper name can be used referentially.

In this chapter we only consider the choice between referential expressions derived from mutual information. We will focus on expressions that serve as answers to *wh*-questions because these emphasize the act of referring in the expressions. But the effect can also be found in other cases of referential use. In particular, the pattern observed for (1) will also be relevant in any other cases of referring to the same individual during the conversation between the same interlocutors. For example, the guest may ask *Is Giacomo coming?* Or the hostess might tell him that *Giacomo has a new boyfriend*. In these cases, however, there may be other issues of appropriateness, i.e. whether the information asserted in the sentence is appropriate. Focusing on short answers allows abstracting from such issues.

1.2. *Referring to locations in English (the trip example)*

Our second example involves reference to locations – unlike the previous example where we only considered proper names and definite descriptions with a contentful head, here we will consider, in turn, all the four types of expressions – proper names, definite descriptions with a contentful head, definite descriptions with a "bleached" head and standard *where* FRs.

1.2.1. Proper names vs. definite descriptions

We start by reproducing the contrast between a proper name and a definite description with a contentful head. Consider the two answers to the question in (3).

(3) Q Where did Dan go on his last vacation?

A1 To London.
A2 To the capital of England.

In a context where the common ground contains the information that London is the capital of England and the speaker wishes just to pick out the relevant entity (as opposed to picking it out by virtue of its having a certain property, such as, for example, being a capital), only the proper name is appropriate – answering with the description is odd. That is, like in the previous example, we see that when used referentially a proper name is preferred to a definite description when the former is available in the common ground – this example might be more intuitive than the one we saw in section 1.1. because it is based on knowledge which is widely shared.

Compare the example in (3) where the equi-extensionality is likely to be mutual information to (4) where the equi-extensionality is less likely to be mutual.

(4) Q Where did Dan go on his last vacation?

A1 To Windhoek.
A2 To the capital of Namibia.

Even if the speaker knows that Windhoek is the capital of Namibia (maybe because her parents are from there), it may be reasonable for her to assume that the addressee doesn't – because this information is not widely available, unlike the equi-extensionality in (3). In this case, the speaker will use the description to refer to the location. Note, importantly, that although this is a reasonable assumption, it is just an assumption and, as such, might be wrong. In section 2.4 we discuss what happens when the speaker makes

the wrong assumptions about the state of the common ground, or, to use Stalnaker's (1978) term, what happens when the context is defective. In the rest of this section, we only discuss examples in which the information is indeed mutual and both interlocutors are aware of this, i.e. we only deal with non-defective contexts.

1.2.2. Definite descriptions with contentful vs. "bleached" nouns

Our next step is to investigate preferences inside the category of definite descriptions. The distinction we make is between a description with a contentful head, like *capital*, and a description with a "bleached" head, like *place*. In the context of reference to locations, a "bleached" noun (*place* or *location*) only encodes the information that the entity denoted is a location, and does not encode additional information about the nature of this location, e.g. whether this location is a building, a city, a continent, or a planet. More generally, a "bleached" noun only encodes the general type of the entity.

We continue to consider the trip example – the two definite descriptions in (5) may in certain contexts be equi-extensional. Let us assume a context in which the common ground contains information that allows constructing these two expressions to refer to the same location, i.e. both interlocutors knows that Dan went to the capital of England last year.

(5) Q Where did Dan go on his vacation?

A1 To the capital of England.
A2 To the place he went last year.

When used referentially, only A1 is an appropriate answer. The "bleached" description sounds like the speaker is being mysterious on purpose, or maybe the speaker wishes to imply that Dan is a boring person that goes to the same place every year. Crucially, the latter would be an attributive use and hence not within the scope of our discussion. That

is, when there is information in the common ground that specifies the nature of the entity beyond its most basic classification, it must be encoded in the expression used to refer to this entity.

1.2.3. Definite descriptions vs. free relatives

Our last distinction concerns descriptions with a lexical head noun vs. descriptions that lack a lexical head noun, such as standard FRs. As pointed out by Caponigro (2003), crosslinguistically FRs introduced by a simple *wh*-word do not allow for a lexical head noun[1] – in this chapter we discuss FRs in English, Italian and Hebrew. It should be noted that this is not true for all FRs – some *–ever* FRs do allow for lexical head nouns; we come back to this in section 3.4.

To compare a referential use of a "bleached" description and a FR, we need to construct a context in which there is a specific location in mind and yet it is not known what kind of place it is – this is not a very common situation. Still, consider a context where you enter the office I share with Dan and see photos scattered on his desk – in the photos we see a beach, a hotel room, and some surrounding buildings – they all characterize a certain location. Crucially, neither of us knows what this place it and where it is located, but I happen to remember these images from his last year's trip. In this context, there is a specific referent in mind, i.e. the location we see in the photos. When answering the question in (6), the bleached description is an appropriate answer –

[1] Internally-headed relatives do not follow this generalization. The languages I focus on in this dissertation do not have internally-headed relatives.

nevertheless, the description still sounds vague[2].

(6) Q Where did Dan go on his vacation?

A1 To the place he went last year.
A2 Where he went last year.

What context would allow using the FR referentially? This would have to be a context in which even a bleached description is not available. Note that the FR differs from the bleached description in that it is not specified for number, i.e. while the bleached description refers to a single location and cannot refer to a group of places (this would require a plural head noun), the FR can refer to either. Accordingly, the FR will be an appropriate answer in a context where it is not known whether Dan went to one place or to a number of places – it is only known that he's taking the same trip as last year.

The distinction between bleached descriptions and FRs is different from the distinctions we saw earlier because there is a logical relation between the descriptions – in the other cases the fact that the two descriptions can refer to the same entity has to be provided by the context, whereas here the bleached description entails the FR.

1.2.4. Summary of the trip example

The trip example shows the following relative preference for the choice of referring expressions: proper names are preferred to definite descriptions with a contentful head

---

[2] Unlike the other distinctions we discuss, this one is not shared by all speakers. The lack of distinction is particular to reference to locations – all speakers seem to have a preference for "bleached" descriptions over FRs when referring to objects – see section 1.4 (the marginality of *who* FRs does not allow testing this distinction in reference to individuals in English, but the preference is found in Italian – see section 1.3). This may be due to the fact that locations are harder to individuate than objects (and people) because of their non discrete nature. In particular, separate locations may be physically contained within each other, e.g. *Rochester* is taken to be a separate location than *New York state* even though it is part of it.

noun, which are preferred to definite descriptions with a bleached head noun which, in turn, are preferred to FRs.

An important question is whether the distinction between the different expressions is due to their linguistic form or rather due to the information encoded in them. While in most cases there is a correlation between the two, there are some cases where expressions that differ in form encode the same information. In particular, although FRs are not grammatically marked for number (because the *wh*-words that introduce them are not marked for number); some FRs denote singular or plural entities because of world knowledge. An example of a singular FR is given in (7): we know that people usually live in one place, therefore A2 is singular. If the preference originates from the form of the expressions, we expect the bleached description to be preferred nonetheless. If, however, it is the information encoded in the expressions that is relevant, no preference is expected. Indeed, when considering the two as an answer to our *where* question, no preference is observed.

(7) Q Where did Dan go on his vacation?

A1 To the place where his mom lives.
A2 Where his mom lives.

This example suggests that the source of the distinction between the different referring expressions arises from the information encoded in the expressions and not from their linguistic form. If this is indeed the case, the four-way classification proposed here is just a by-product of the kinds of information that are normally encoded using the different forms. In what follows, I will continue to use this four-way classification, keeping in mind that what we really need is a fine grained measure of the information encoded in the different expressions.

Independent of how the information can be measured, the trip example shows that if interlocutors are familiar with certain aspects of an entity, these aspects must be encoded in an expression used to refer to that entity. The preferences are summarized in (8): the speaker should use whichever information is available in the following order.

(8) (i) A proper name.
(ii) The sortal concept to which the entity belongs, i.e. the common noun which characterizes the nature of the entity (the term 'sortal concept is adopted from Gupta 1980).
(iii) Number.

Note that even the least specific expressions (I am using specific here in some intuitive sense) encodes information about the nature of the entity. For example, although FRs do not encode any of the kinds of information mentioned in (8), the *wh*-words the introduce them do encode some general information about the entity: *where* FRs denote locations, *who* FRs denote people, *how* FRs denote manners and *what* FRs denote object, properties or propositions[3].

1.3. *Referring to individuals in Italian (the Lasagna example II)*

In this section we now go back to the lasagna example and examine other expressions that refer to individuals. This example will be presented in Italian for two reasons. First, since nouns in Italian are all marked for gender, this example will allow us to consider gender information. Second, FRs introduced by *who*, which are rather restricted in English (see again chapter 2, example 12), are very natural in Italian.

---

[3] Note that *what* FRs can denote more than one type of entity, contrasting with *who* FRs which always denote individuals and *where* FR which always denote locations. The entity denoted by a *what* FR can be an object (*what I bought*), a property (*what Dan is*), or a proposition (*what I said*) – the kind of the entity is therefore determined only by the property in the relative clause.

1.3.1. Proper names vs. definite descriptions

First, consider (9) which is the Italian counterpart of (1).

(9) Q Chi ha portato le lasagne?
who has brought the lasagna
'Who brought the lasagna?'

A1 Giacomo
A2 Il vicino della porta accanto
the neighbor of-the door next
'The next-door neighbor.'

Like in English, in a context in which it is mutual information that the proper name and description in (9) are equi-extensional, the proper name is the appropriate answer and using the description is odd. This further supports our generalization that if the proper name of a certain entity is known, it must be used in referring to that entity.

1.3.2. Definite descriptions with contentful vs. "bleached" nouns

Our next example compares the appropriateness of a definite description with the contentful noun *vicino* 'neighbor' with a description with a bleached noun – *ragazzo* 'guy'. Note that both these descriptions are marked for gender – they are masculine in accordance with the gender of the head noun.

(10) Q Chi ha portato le lasagne?
who has brought the lasagna
'Who brought the lasagna?'

A1 Il vicino (della porta accanto)
the neighbor of-the door next
'The (next-door) neighbor.'

A2 Il ragazzo con il cappello
the guy with the hat
'The guy with the hat.'

Consider a context in which the hostess of the party is talking to one of the guests, while the individual who is the referent of the answer is also in the room. The hostess, i.e. the

speaker who picks the referring expression, knows her neighbors, so the crucial thing is what the guest, i.e. the addressee, knows (or rather, what the speaker assumes he knows). If the addressee also knows that this is the neighbor, the description with the contentful noun is preferred. Using the bleached description in the same context is odd – it gives the feeling of less familiarity with the individual. This shows that if a sortal concept is available, it must be expressed in a referring expression even if it is not required to pick out the entity, i.e. if successful reference can be established without it.

1.3.3. Definite descriptions with "bleached" nouns vs. free relatives

Next, we compare the bleached description with a *who* FR which is unspecified for both number and gender (for the *where* FRs we saw in section 1.2, gender was irrelevant). This contrasts with headed descriptions – both contentful and bleached – which are grammatically marked as singular or plural and also as masculine or feminine.

(11) Q Chi ha portato le lasagne?
who has brought the lasagna
'Who brought the lasagna?'

A1 Il ragazzo con il cappello
the guy with the hat
'The guy with the hat.'

A2 Chi indossa il cappello
who wears the hat
Roughly: 'Who is wearing the hat.'

Consider again the context where the interlocutors are in the same room with this person: the mutual information is available by physical co-presence (see Clark & Marshall 1981): they see a guy wearing a hat. In this context, the bleached description is appropriate while the FR is odd – it gives the impression that the speaker is not being "specific enough" (to quote one of my consultants). This example shows that when gender information is

mutual, it must be expressed in the referring expression, even if successful reference can be established without using this type of information.

When would the FR be appropriate? Since wearing a hat is necessarily a property of one individual, number information cannot distinguish the FR from the bleached description – with this property both expressions are singular. However, the two descriptions differ in gender information – the bleached description is marked masculine while the FR is unspecified for gender and can denote a masculine or feminine individual. It is possible to think of a context where gender information will not be available – if the interlocutors see the individual from the back and can't tell if it's a man or a woman (the person is dressed in unisex clothes). In this case, there isn't sufficient information to use a headed description and the FR would be licensed.

We consider another example of a choice between a bleached description and a FR in which the property does not necessitate physical co-presence.

(12) Q Chi ha portato le lasagne?
who has brought the lasagna
'Who brought the lasagna?'

A1 Il ragazzo che è arrivato durante il temporale
the guy that is arrived during the storm
'The guy who arrived during the storm.'

A2 Chi è arrivato durante il temporale
who is arrived during the storm
Roughly: 'Who arrived during the storm.'

In a context where the interlocutors have seen an individual come in during the storm, we get the same pattern where the headed description is appropriate and the FR is odd. The only context which would license the FR is such that would not allow using the bleached description. This would be a context where there is no gender or number information available about the individual. In particular, if the interlocutors didn't see whoever came

in during the storm – they only heard the door, and the lasagna was a bit wet on the top, then there just isn't the information to use A1 (or a feminine or a plural alternative). This again shows that when certain types of information are available in the common ground, i.e. number and gender, they must be expressed in the referring expression, even successful reference can be established without them.

1.3.4. Summary of the lasagna example

The Italian version of the lasagna example again raises the question of whether it is the form of the linguistic expression that underlies the preference or the information encoded in the different expressions. Like in the English trip example, we consider a case where the information encoded in the FR is equivalent to what is encoded in a bleached description. Consider (13) in a context of a running competition we were having in the yard during the potluck. The information encoded in the bleached description is the same as what is encoded in the FR: both expressions denote a singular individual (because only one person can be first) and are unspecified for gender (*persona* can refer to both women and men).

(13) Q Chi hanno premiato?
who have rewarded
'Who have they rewarded?'

A1 #?La persona che è arrivato per primo
the person what is arrived for first
'The person who arrived first.'

A2 Chi è arrivato per primo
who is arrived for first
Roughly: 'Who arrived first.'

Unlike the English example in (7) where both expressions were acceptable, in Italian the pattern of preference reverses: the bleached description "sounds redundant" and is hence less preferred. This might be because the FR is shorter.

The conclusion, then, is the same as for English: the pattern discussed here concerns the information encoded in the referring expressions and not the form of these expressions. When two expressions contain the same information, the choice is determined by other factors, one of which would be the length of the expression.

1.4. *Referring to objects in English (the soufflé example I)*

Our next example concerns reference to objects. Here, proper names are not relevant because, unlike individuals and locations, objects do not have names.

1.4.1. Definite descriptions with contentful vs. "bleached" nouns

We start by comparing a contentful description with a bleached description. We continue to use a potluck context and consider a situation where a very picky guest is trying to figure out what to eat. A second guest who is already eating wants to recommend the soufflé which is in a blue bowl – note that the bowl is physically co-present. Consider the question with the two possible answers in (14).

(14) Q What's worth eating?

A1 The soufflé.
A2 The thing in the blue bowl.

Our context of interest is such that the speaker knows what soufflé is, can identify the contents of the blue bowl as such, and, crucially, also assumes that this knowledge is mutual. In this context, the description with the contentful head is appropriate – answering with the bleached description is odd; while this may lead to successful

reference, it seems like the speaker is being vague on purpose. In fact, if the speaker used the bleached description in such a context, the addressee may overtly protest by saying: "I know what soufflé is!" (we come back to such mismatch in section 2.4). That is, using the bleached description implies that the fact that the entity being referred to belongs to the sortal concept *soufflé* is not mutual information.

One could argue that given the potluck context, it is mutual information that whatever is in the bowl is food and, therefore, there is a contentful noun available for the entity in question, i.e. *food*, and it should be preferred to a bleached noun like *thing*. However, I argue that this head noun is not appropriate for a different reason. In particular, Brown (1965) proposes that although there are different lexical nouns which are true of the same object, some nouns are most commonly used when talking about these object. For example, an entity that belongs to the class of *fruit*, *apple* and *Granny Smith* will most frequently be referred to as *an apple*. Brown suggests that this is because the information expressed in BASIC LEVEL TERMS, such as *apple*, is in most cases the most important or relevant for the purposes of communication – more so than the information that this entity belongs to the class of fruit, which is the SUPERORDINATE LEVEL TERM, or to the class of Granny Smiths, which is the SUBORDINATE LEVEL TERM.

Cruse (1977) adopts Brown's distinction, but proposes that the underlying motivation for using one noun over the other is different. In particular, he views the nouns as differing on their level of specificity and proposes that the speaker chooses the level of specificity depending what contrast is relevant in the particular context, e.g. a contrast between apples and pears, or a contrast between different varieties of apples. Independent of which explanation we adopt, using a superordinate level term, such as *food* or *dish*, in

(14) is dispreferred for a different reason, hence should not be considered as a potential referring expression when discussing levels of familiarity.

1.4.2. Definite descriptions with "bleached" nouns vs. free relatives

Next, we compare the bleached description with a FR introduced by *what*, as in (15). Importantly, we are considering which of these expressions is a better referring expression for the same referent we considered in the previous section, i.e. the soufflé in the blue bowl, but we change the context such that it is not mutual information that this entity is *soufflé* and hence this noun cannot be used. In this context, the bleached description is preferred to the FR.

(15) Q What's worth eating?

A1 The thing in the blue bowl.
A2 What's in the blue bowl.

Unlike with previous examples, here it is hard to come up with a context which would license using the FR felicitously. This would require a context in which no bleached noun, such as *thing* or *stuff*, is available, which is not possible given that we know that it is "in the blue bowl"; even if we can't see the contents of the bowl and we don't know if it is count or mass, we can make a reliable assumption based on the context: something in a bowl will be referred to as *stuff*.

Let us then consider a property which does not require physical co-presence, as in (16). We first consider the choice of referring expressions in a context in which there is a specific referent in mind. For example, here I know you had your mom over and that you always cook one very special dish for her (but it's a different one every time, or I don't remember what it's called, so the sortal concept is not mutual information). In answering the question in (16), I am trying to refer to this dish. In this context, like in the soufflé

example in (15), the bleached description is preferred.

(16) Q What would you like to eat?

A1 The thing you cooked yesterday.
A2 What you cooked yesterday.

But consider the same example out of the blue, i.e. when there isn't a specific dish in mind that the speaker wishes to pick out – it is only mutual that the addressee cooked something yesterday, say because he always cooks on Fridays. In this context the nature of the referent is vague: it might be a cake (i.e. a *thing*), it might be jambalaya (i.e. *stuff*), or it might be both which would require using a plural description. In this context the speaker will use a FR because there isn't enough information about the referent to allow using a head noun. This contrasts with a context where a specific (singular) referent is mutual knowledge in which this information must be encoded in the referring expression.

1.5. *Referring to objects in Hebrew (the soufflé example II)*

Hebrew, unlike English, does not have a word which is the equivalent of *stuff* (the use of the equivalent of *thing* is much more restricted), so no bleached description can be constructed for the soufflé example – we can only compare a contentful description with a FR. Consider again the potluck context – the interlocutors and the blue bowl with its contents are physically co-present and it is mutual knowledge that its content is soufflé.

(18) Q ma keday le-exol?
what worth to-eat
'What's worth eating?'

A1 et ha-sufle
Acc the- soufflé
'The soufflé'

A2 et ma še-ba-ke'ara ha-kxula
Acc what that-in-the-bowl the-blue
'What's in the blue bowl.'

Even though each of the expression can uniquely pick out the soufflé, only the contentful description is an appropriate answer in this context – using the FR implies that the noun *sufle* 'soufflé' is not mutual knowledge.

The Hebrew example differs from the English one in the context that allows using the FR. In particular, because of the difference in availability of a bleached description, a context in which it is not mutual information that the contents of the blue bowl is soufflé is enough to license the FR in Hebrew but not in English. From a different perspective, the FR in Hebrew will imply the lack of a sortal concept, while the FR in English also implies the lack of more basic information about the nature of the entity (and/or number information). This contrast shows that evaluating referring expressions depends on what other choices are available to interlocutors (which in turn depends on the specific language used) and not just on the information encoded in the expression considered.

1.6. *Summary*

This section presented data showing that when the common ground contains information that allows constructing more than one referring expression for a certain entity, some expressions are preferred to others. In particular, proper names are preferred to headed descriptions with a contentful noun, which are preferred to bleached descriptions, which, in turn, are all preferred to FRs.

An important question is whether the hierarchy comes from the linguistic form per se, or whether the information conveyed in each of these forms tends to line up with these constructions. I have tried to address this question by considering expressions from

different categories that contain the same information. In particular, we have compared bleached descriptions with FRs that are marked for number (or gender) by virtue of the property expressed. For these cases, we saw that the pattern of preference changes – see again examples (7) and (13) – and hence concluded that it is the information itself and not the form of the expression that is relevant for our pattern. The pattern is summarized in (19), revised from (8).

(19) Use the following information when referring to an entity:
(i) A proper name, if mutually known.
(ii) The sortal concept (Gupta 1980) to which the entity belongs, i.e. the basic level term, if mutually known.
(iii) Number and/or gender, if mutually known.

Recall that even FRs, which we saw to be the least preferred, encode some basic information about the nature of the entity: *where* FRs denote locations, *who* FRs denote people, and *what* FRs denote objects, properties or propositions.

I propose that the referring expression used reflects the level of familiarity of the interlocutors with the entity in question. The next section proposes how to analyze this pattern.

2. A Scalar Phenomenon

This section argues that the preference phenomenon we saw in the previous section is an instance of scalar implicatures. In particular, in choosing among possible referring expressions to refer to an entity the speaker must use the expression that best reflects the familiarity of the interlocutors with the entity in order to comply with Grice's Maxim of Quantity which requires the speaker to provide all the information available to her.

The section is organized as follows. The first part argues that we are dealing with a scalar phenomenon: section 2.1. briefly introduces the phenomenon of scalar implicatures

and evaluates how the pattern discussed here fits in. Section 2.2. and 2.3. discuss two apparent counter-examples and show how they fall within the pattern once Gricean principles are taken into account. Section 2.4. presents some ideas concerning how the scale is formed – unfortunately, I am not able to provide a definition of familiarity at this point. Finally, in section 2.5. I address the conversational significance of the phenomenon – we will see that the interlocutors use the choice of referring expressions to make sure they are "on the same page" with respect to their knowledge about the entities.

2.1. *Scalar implicatures*

Grice's Maxim of Quantity (Grice 1975) states that the cooperative speaker must provide all the information available to her. It is comprised of two parts: Q1 requires her to provide all the information, and Q2 requires providing no more than what is really needed.

(20) Q1 Make your contribution as informative as required.
Q2 Do not make your contribution more informative than is required.

Consider a standard scalar implicature, such as the one arising due to the scale of quantifiers *all* > *some*.

(21) a. Some kids eat broccoli.
b. All kids eat broccoli.

If the speaker knows that a subset of kids eat broccoli (context α), only (21a) is a true utterance. If, alternatively, she knows that all kids eat broccoli (context β), she can truthfully utter both (21a) and (21b), but only (21b) would give all the information available to her. The choice between (21a) and (21b) in this case is governed by Q1 which requires the speaker to provide <u>all</u> the information available to her.

We can also view this from the addressee's perspective. In particular, if the speaker utters (21b) – the situation must be such that all kids eat broccoli (context α). If the speaker utters (21a) – it necessarily means that a subset of the kids eat broccoli but no information is given about the rest of the kids – they may or may not eat broccoli. Logically, the situation is ambiguous between all kids eating broccoli (context α) and only a subset of kids eating broccoli (context β). However, since speakers are assumed to be cooperative unless there is a reason to think otherwise, the addressee would assume that all the information available to the speaker was conveyed by (21a), and hence conclude that not all kids eat broccoli, i.e. that (21b) is false.

How does this compare with the phenomenon discussed in the previous section? Consider again the lasagna example, repeated as (22).

(22) Q Who brought the lasagna?

A1 Giacomo.
A2 The next-door neighbor.

When both expressions are available, i.e. there is information in the common ground that allows constructing either expression, we find the same pattern as with standard scalar implicatures: the stronger term must be used, which here is the proper name. But unlike with standard scales which are ordered by entailment, here there is no logical relation between the terms on the scale, and thus each expression can be available independently. That is, while for the quantifier scale there can be no situation where (21b) is true but (21a) is false, such a context can be found for (22) – namely any context where Giacomo is not at all the next-door neighbor.

Turning to the addressee's perspective, we also find similarities and differences. The similarity is that when the weaker term on the scale, i.e. the description, is used, it is

implicated that the speaker was not in a position to use a stronger term. The difference lies in why the speaker cannot use the stronger term: for standard logically ordered scales this is because the sentence containing the stronger term is false, i.e. the implicature concerns what is asserted. For our scale here, on the other hand, if the speaker does not use a stronger term, the implicature is that the common ground does not contain information that would allow constructing a stronger term for that entity, i.e. the implicature concerns what is presupposed[4]. A second difference lies in what is implicated when the stronger term is used: with standard scales, the sentence containing the weaker term is entailed, but with our scales nothing is implicated – the weaker term may or may not be available in the common ground. This difference is also a consequence of the lack of logical relations in the proposed scale.

The similarities can be attributed to the fact that both phenomena are implicatures involving scales: a weaker term negates all stronger terms on a scale. The differences, in turn, can be attributed to the different nature of the two scales: standard scales are ordered by entailment so they are available independent of a particular context, while the scales here are specific to a certain context. As a result, a standard scalar implicature can be computed independent of the context, but the implicatures here can only be computed relative to a specific context: a certain choice of referring expressions constitutes a signal about the state of a specific common ground.

Note that this type of scalar implicatures are similar to the implicatures discussed by Gundel, Hedberg & Zacharski (1993) in the context of their analysis of the choice of

---

[4] An explicit distinction between information about the world and discourse information is made by Groenendijk, Stokhof & Veltman (1996).

anaphoric expressions based on the giveness status of the referent. We come back to this analysis in section 3.1. when discussing the relevance of other types of expressions for our pattern.

2.2. *Apparent counter-example I: flouting the Maxim of Quantity*

In all the examples we saw so far the speaker was cooperative and used the strongest term on the contextually-determined scale of referring expressions. But there are cases where the speaker can felicitously use the weaker term on the scale. Consider the examples in (23-24) where the interlocutors are parents.

(23) Q Who stained the sofa?
A Your daughter.

(24) Q Who won the prize?
A My daughter.

Given the identity of the interlocutors, it is clear that the information about the name of the individual is mutual and it is equi-extensional with the description. Nonetheless, using the description is appropriate. This contrasts with the examples we saw in section 1. Crucially, though, using the description conveys another message in addition to picking out the entity. In (23) the extra message is attributing the responsibility of the daughter's behavior to the other parent and in (24) it is attributing the responsibility of the daughter's achievement to themselves.

While not as easy to see, the same effect can be created for (at least some of) the examples we saw earlier. For instance, consider again the lasagna example in a context where the common ground contains information that allows associating both the proper name *Giacomo* and the description *the next-door neighbor* with the same entity. Assuming the interlocutors are confident that this information is indeed mutual and the

speaker still refers to the individual in question using the description, the addressee might see this as a conscious choice that was made in order to convey an extra message – maybe that she does not like the neighbor and hence avoided his name.

These examples may seem prima facie as counter-examples to the pattern presented in section 1; a description is felicitously used when a proper name is mutually available. However, the emergence of the extra message makes the two contexts crucially different. I argue that it is the communication of the extra message that licenses the use of the description – if such a message cannot be easily recovered as was the case in the examples presented in section 1, using the description as an answer is odd. This is another piece of evidence that using the strongest term on the scale is the unmarked option.

What is the extra message? I propose that this is an implicature that is caused by a deliberate flouting of the cooperative principle. In particular, the speaker deliberately uses a weaker term on the scale in order to flout Grice's Maxim of Quantity (in particular, Q1). Recognizing the flouting, the addressee tries to find a reason for it by trying to recover the message the speaker intended to convey by the flouting. In (23), using a description with the possessive *your* can be interpreted as taking responsibility off the speaker and attributing it to the addressee. This is supported by the fact that a negative property is mentioned. In (24), the message is the opposite – using the possessive *my* assigns the responsibility to the speaker. Finally, in the lasagna example (see again example 1), avoiding using the name of the neighbor can be interpreted as a signal of dislike – in the case presented here it was not strongly supported by the content of the utterance and thus was harder to recover.

Note that the actual content of the flouting implicature varies between the different examples – it is not encoded in the context or in the utterance itself and rather has to be constructed based on the combination of these with world knowledge. Note also that I am not trying to explain how this message is constructed – what is relevant for my purposes here is not how this content is computed, but rather the fact that flouting implicatures emerge at all.

2.3. *Apparent counter-example II: special nouns: 'mother' and 'father'*

Another deviation from the pattern presented in section 1 concerns two specific lexical nouns – *mother* and *father*. These kinship terms can be used even when the proper name of the relevant individual is available in the common ground. In fact, using the proper name in these contexts is odd.

Consider, for instance, a context of me entering my office, asking the question in (25-26). Let us assume that my officemate is well acquainted with my parents, in which case there is no doubt that she knows their names. Although both the proper names and the descriptions are available in the common ground, the description is the only felicitous answer – using the proper names is odd[5].

(25) Q Who called? A1 Your father
A2 Moshe / Mr. Heller.

(26) Q Who called? A1 Your mother.
A2 Ariella / Mrs. Heller

[5] Note that the more formal versions *Mr. Heller* and *Mrs. Heller* might be available even without acquaintance. This, however, is irrelevant for our purpose here, because the point if that the description is preferred. Note, incidentally, that deducing the parents' names based on the child's might be wrong – people do not always have the same last name as their parents.

This pattern is the mirror image of what we saw in section 1 for other nouns that could have not been used when a proper name was also available. Interestingly, this pattern is unique to *mother* and *father* and it is not found with other kinship terms, whether they express a unique relation, such as *husband*, or a non unique relation, such as *sister*.

I propose that this special pattern reflects the convention in our culture that one does not use proper names to refer to one's parents, but instead uses the terms *mother* and *father* (and shorter versions thereof). This convention renders it inappropriate to use proper names if parenthood relationship is relevant to one of the interlocutors – in these cases this convention overrides the factor of discriminability.

2.4. *Familiarity and discriminability*

Throughout the chapter, I have proposed that the referring expression used to refer to a certain entity has to reflect how familiar interlocutors are with that entity. Intuitively, knowing a proper name of an individual or a location is a signal that one is more familiar with that entity than if one knows its sortal information. Sortal information, in turn, reflects better familiarity with the entity than knowing the general type of the entity, i.e. if it is an object, a person or a location.

But why is an expression that reflects a higher level of familiarity preferred to an expression that reflects a lower level of familiarity? I propose that this is because the former provides a better identification of the entity denoted. That is, the more familiar interlocutors are with an entity, the more information they have that allows them to identify it. One possibility for quantifying identification is by comparing how well different expressions allows interlocutors to discriminate the entity from other entities.

It is important to point out that identification is different from the notion of reference (or picking out). In particular, establishing successful reference requires the speaker to provide enough information to allow picking out the entity in the specific context; the information required depends on the specific context. Identification, by contrast, depends on the information available about a certain entity that would allow it to be identified across contexts.

For example, if we know Dan was born in a city, we can refer to it using a description like *the city Dan was born in* – this will express sufficient information to pick out the place. But how well does this information identify the place? This depends on how well it discriminates it from other places, i.e. which places are necessarily distinct such that they cannot be collapsed with this place when information is added to the common ground. The place in question can potentially be collapsed with entities denoted by *London*, *the city in which I was born* or *the place where Dan and I met*, but it is necessarily distinct from *Namibia* (which is a country, not a city) and *the island Dan lives on*. The idea is that the more familiar we are with an entity, the better we identify it in that it is better discriminated from other entities.

Unfortunately, at this point I am unable to provide a formal definition of how identification is measured. I hope to take up this issue in future work. Thus, for now, we will continue to assume that a measure can be established that will allow ranking proper names over descriptions with a contentful head, over descriptions with a bleached head, and over descriptions that lack a head.

2.5. *The effect of the phenomenon: maintaining a non-defective context*

The main goal of a referential use of an expression is to pick out a certain entity in order to say something about it – this can be equally achieved by any expression that leads to successful reference. The phenomenon presented in this chapter suggests that the choice of referring expressions has an additional role.

Let us take a step back and consider the notion of common ground. While it is common practice in the literature to simplify matters and assume a conversation has a single common ground (e.g. Heim 1982, Kadmon 2001), we should keep in mind that interlocutors do not in fact have access to that such shared common ground. Rather, each interlocutor has her own representation of the common ground. This is explicitly stated in Stalnaker (1978), who defines the notion of common ground as "the presuppositions of a speaker". Stalnaker distinguishes a non-defective context in which the common grounds are the same from a defective context in which there are mismatches between the interlocutors' common grounds, and claims that the latter will lead to failure of communication and will thus have to be fixed.

Stalnaker does not directly discuss how interlocutors go about making sure that their common grounds are the same, but he does expresses his view in passing when discussing defective contexts:

> "Since communication is the point of the enterprise, everyone will have a motive to try to keep the presuppositions the same. And because in the course of a conversation many clues are dropped about what is presupposed, participants will normally be able to tell that divergences exist if they do." (p. 322)

What are these clues? I propose that the phenomenon presented in this chapter is one such clue. In particular, when the speaker chooses to use one referring expression instead of another, she not only picks out the entity she wishes to say something about, but also

signals to the addressee what she assumes the state of the common ground to be. When the addressee, in turn, accepts the speaker's choice of referring expression, he also accepts a certain state of the common ground. That is, a referring expression serves as a means for the interlocutors to try to keep their common grounds the same.

This is supported by the fact that if the addressee does not accept the suggested state of the common ground, he may object to the speaker's choice of referring expression. Consider, for example, another version of the lasagna example where the addressee knows Giacomo is the next-door neighbor, but the speaker mistakenly assumes that he does not know the name of that individual. In such a defective context, the speaker will use the description to refer to that individual, reflecting her assumed state of the common ground. The addressee, recognizing that the context is defective, will try to bring the two common grounds together by discussing the issue – he may say something along the lines of "You mean Giacomo?!? I've met him numerous times!" or "Didn't you know he's my cousin?" or "You forget you told me a lot about him?". This objection to the state of the common ground proposed by the speaker will lead her to correct her common ground. When no objection is raised, the speaker will assume that the addressee has accepted the state of the common ground she has put forward.

In sum, we see that choosing among possible referring expressions is not arbitrary. It serves a secondary purpose in addition to picking out the individual. In particular, it serves as a clue about the assumed state of the common ground and helps interlocutors in their goal of keeping their common grounds the same.

3. Other Expressions

The expressions discussed so far have been proper names, definite descriptions and FRs, but there are of course other expressions in natural language that denote entities. In this section, we consider how indefinites (section 3.2), pronouns (section 3.3) and *–ever* FRs (section 3.4) fit into the picture. But, first, I present Gundel, Hedberg & Zacharski's (1993) analysis of the choice of anaphoric expressions (section 3.1). The reason for presenting this analysis is twofold. First, Gundel et al. discuss a different dimension along which referring expressions are chosen – the cognitive status of the referent – which will be relevant when discussing indefinites and pronouns. Second, Gundel et al.'s analysis shares some properties with the analysis presented here – both deal with a scalar phenomenon that concerns the state of the discourse rather than what is asserted.

3.1. *Gundel, Hedberg & Zacharski's (1993) scale of anaphoric expressions*

Gundel, Hedberg & Zacharski (1993) use a scale to analyze the distribution of a different range of referring expressions – indefinites, definites, demonstratives and pronouns. The scale they propose orders cognitive statuses, which they refer to as "the Giveness Hierarchy" adopting a common term in the literature on the choice of anaphoric expression. Their proposed scale is given in (27).

(27) in focus > activated > familiar > uniquely identifiable > referential > type identifiable

Crucially, each status on the scale entails all the lower statuses. That is, if a referent is *uniquely identifiable*, it is also *referential* and *type identifiable* (note that Gundel et al. use the term *type* for what we call here *sort*, i.e. the default common noun to which the entity belongs).

Each of these cognitive statuses is a necessary and a sufficient condition for using a certain type of anaphoric expressions. The correspondence is illustrated in (28).

(28)

in focus >	**activated >**	**familiar >**	**uniquely identifiable >**	**referential >**	**type identifiable**
It	*that* *this* *this* N	*that* N	*the* N	(indefinite) *this* N	*a* N

If the cognitive status of a referent is *uniquely identifiable*, it fulfils the necessary and sufficient conditions for using an expression of the form *the N*. But this referent also fulfils the conditions for all the entailed cognitive statuses – *referential* and *type identifiable*, and hence is also compatible with the expressions that correspond to these statuses, i.e. *a N* and *this N*. The prediction then is that in natural discourse the different forms will be distributed across more than one cognitive status. In a crosslinguistic corpus study, Gundel et al. indeed found that the same form was used for more than one cognitive status. However, the distribution of the forms is more restricted than what is predicted by the Giveness Hierarchy alone. Gundel et al., propose that the choice of anaphoric expressions is governed by an interaction of the Giveness Hierarchy with Grice's Maxim of Quantity. For ease of reference, the two parts of the Maxim are repeated in (29).

(29) Q1 Make your contribution as informative as required.
Q2 Do not make your contribution more informative than is required.

Gundel et al. show that the interaction between their proposed Giveness Hierarchy and the cooperative principle can account for the distribution of anaphoric expressions they found.

First, indefinite expressions refer to entities that are at most *referential* – there are no indefinites that refer to entities whose status was *uniquely identifiable* (or higher). This

pattern is accounted for by Q1 which requires the cooperative speaker to give all the information available to her, so if the speaker knows that the cognitive status of the referent is *uniquely identifiable*, this information should be encoded in the choice of the anaphoric expression, i.e. the cooperative speaker should use a definite expression.

A second case of interaction of the Giveness Hierarchy with Q1 is in the choice of referring expression for entities whose cognitive status is *in focus*. Since being *in focus* entails being *activated*, we might expect to find entities whose status is *in focus* that are being referred to using a demonstrative pronoun. Such cases are, however, not found. Gundel et al. attribute this to Q1 which requires the cooperative speaker to give all the information. Choosing a referring expression that signals the status *activated* for an entity whose status is in fact *in focus* would violate this part of the Maxim.

The third case of interaction with Grice's Maxim of Quantity involves Q2. Gundel et al. find that definite descriptions are used for entities that have a higher cognitive status than *uniquely identifiable*. In particular, they found definites for cases where the entity is *familiar*, *activated* or *in focus*. This choice, they suggest, accords with Q2. In particular, the uniqueness information conveyed by *the N* suffices to identify the entity. Using an expression that corresponds with a higher cognitive status would require an expression that contains less information about the nature of the entity being referred to – pronouns do not encode such properties – just the number, gender and animacy of the entity (whichever is relevant). In these cases, Gundel et al. claim, encoding sortal information will be preferred because the expression associated with the lower cognitive status can still lead to successful reference.

The interaction between the Giveness Hierarchy and Grice's Maxim of Quantity allows accounting for a complex pattern. In this respect, this analysis differs from other analyses that use hierarchies in which the anaphoric expressions directly correspond to cognitive status, e.g. Ariel (1988).

This analysis shares some properties with the analysis presented here. First, the two are similar in that they take the choice of referring expressions to be a scalar phenomenon. Further, the scales are similar in that they deal with facts about the state of the discourse and not the state of the world, i.e. about what is asserted. Finally, both analyses assume that there are other considerations in choosing a referring expression with which the factor presented interacts. While sharing some concepts, the two analyses differ with respect to the actual phenomenon they deal with and, as a result, with the range of expressions they consider: Gundel et al.'s analysis deals with the salience of entities, which is affected by when and how these entities have been mentioned in the discourse, while my analysis deals with the familiarity of interlocutors with the entities, which is affected by how much information they have about it. A complete analysis of the choice of referring expressions will have to take both dimensions into account[6].

3.2. *Indefinites*

In the previous section we saw that indefinites differ from definites in that the entity they denote does not uniquely fit the property mentioned in the description. But when

---

[6] Comparing Gundel et al.'s analysis and mine seems to imply that Ariel's (1988) analysis of the choice of referring expressions is in the wrong direction. In particular, Ariel's scale contains pronouns and definites, like Gundel et al.'s, but also proper names, like the scale proposed here. The discussion here suggests that choosing between a proper name and a definite description is different from choosing between a description and a pronoun.

abstracting away from uniqueness, which concerns the discourse status of the referent and not its identity, the two kinds of descriptions encode the same information. In particular, like definites, indefinites encode sortal information, number and gender. Accordingly, we expect them to be of the same discriminability as definites.

However, the conditions under which we were examining the choice of referring expressions in this chapter are not appropriate for the use of indefinites. In particular, the contexts we have been considering are such that the speaker wishes to talk about a specific entity in the common ground and chooses the information that will be used to pick it out. If the speaker has a certain entity in mind, that means it is *uniquely identifiable* and, as Gundel et al. have shown, will not be picked out using an indefinite.

But the exclusion of indefinites from the expressions considered here does not depend on adopting Gundel et al.'s analysis. If instead we adopt the influential analysis of Kamp (1981) and Heim (1982) in which indefinites introduce new entities into the common ground, we end up with the same conclusion – such an expression will not be used for an entity already in the common ground. We can also view this from a different perspective: since the entity introduced by an indefinite is not in the common ground, it will not be associated with different types of information. Thus, it is irrelevant to talk about choosing which of these pieces of information will be encoded in the referring expression.

Whichever of these analyses we adopt, indefinites will not be available as referring expressions for specific entities in the common ground, and hence they cannot be compared to other expressions when choosing a referring expression for a specific entity in the common ground. This does not necessarily mean that indefinites are irrelevant for discriminability. We predict indefinites to be of similar discriminability to definites, as

they encode the same kinds of information. In the next chapter we discuss discriminability in specificational sentences, where the discriminability of indefinites will be relevant.

3.3. *Pronouns*

We have mentioned in section 3.1. that, according to Gundel et al. (1993), pronouns encode minimal information about their referents. In particular, a pronoun may encode the number and gender of the entity (or entities) denoted and also whether it is animate. In addition, the entity denoted by a pronoun is already in the common ground. Thus, pronouns should be considered in the context of choosing a referring expression based on discriminability.

The prediction is that definite descriptions will be preferred because in addition to number, gender and animacy information, they also encode the sortal information of the entity. On the other hand, pronouns and definite descriptions differ in the cognitive status they correspond to. In particular, a pronoun requires its referent to have the cognitive status *in focus* – this is a subset of the cases where the cognitive status of the referent is *uniquely identifiable*, the status required for using a definite description. Thus, comparing the two will only be relevant when the entity is *in focus* when both types of expressions are available. In these cases, using a pronoun will satisfy the consideration of giving all the information regarding the cognitive status of the referent whereas using a definite description will satisfy the consideration of giving all the information regarding the familiarity of the interlocutors with the referent. Which one is more important? The data presented by Gundel et al. shows that in most cases the salience factor wins out (87%, 214/246), yet there are cases in which the definite is still preferred (12%, 30/246) – I take

this to show that discriminability in some cases wins out. This pattern is an example for how factors interact in governing the choice of referring expressions.

3.4. *-ever FRs*

The last kind of expression we discuss is *-ever* FRs. We will see that, under some analyses, *-ever* FRs are attributive-only expressions which makes them irrelevant to the choice of referential expressions based on familiarity. However, even if we do not adopt this analysis, we will see that the contexts that license *–ever* FR are such that neither proper names nor descriptions can be used and hence the discriminability of *-ever* FRs cannot be compared with that of proper names and definite descriptions.

-ever FRs are special in that they introduce indeterminacy with respect to the identity of the referent. More specifically, when a speaker uses an *–ever* FR, she signals that she doesn't know – or doesn't care – what the identity of the referent is. Consider *–ever* FRs as answers to the questions we used for our examples in section 1.

(30) Q Who brought the lasagna?
A Whoever came during the storm.

(31) Q Where did Dan go for Christmas?
A Wherever he went last year.

(32) Q What's worth eating?
A Whatever is in the blue bowl.

The *whoever* FR in (30) means that the speaker does not know (or doesn't care) who it is that brought the lasagna. This FR will not be licensed if she knows that it was Giacomo or the next-door neighbor. The *wherever* FR in (31) means the speaker only knows that Dan went to the same location as last year, but she doesn't know, or doesn't care, where this place is; the FR cannot be used if the speaker knows that this place is London, the capital of Namibia, or an island in the Pacific. The *whatever* FR in (32) means that the

speaker doesn't know, or doesn't care, what the contents of the blue bowl is; if the speaker knows that it is soufflé (or clam chowder), the FR cannot be used.

Although the identity of the referent is not known, *-ever* FRs have nonetheless been argued to denote entities (Jacobson 1995)[7]. But how is the indeterminacy captured? Dayal (1997) proposes that *-ever* FRs are unambiguously attributive, in the sense of Donnellan (1966). In particular, when using an *–ever* FR, the speaker does not have a particular entity in mind she wishes to pick out. Instead, an *–ever* FR picks out an entity by virtue of the property expressed in the FR. Dayal represents attributity by introducing a modal dimension to the denotation of *–ever* FRs and varying the identity of the entity denoted across the epistemic alternatives of the speaker. If *-ever* FRs are unambiguously attributive, they are irrelevant to the pattern discussed here; we saw in section 1.1 that when expressions are used attributively, they are chosen depending on the property they encode and not based on the familiarity of the interlocutors with the entity.

While Dayal's attributive-only analysis for *-ever* FRs was adopted by Quer (1998) (who also extends it to subjunctive FRs in Spanish and Catalan), other researchers, such as von Fintel (2000) and Tredinnick (2005), do not refer to the attributive-referential distinction in their analyses. If *-ever* FRs are not attributive-only, they should be compared with other expressions for discriminability.

---

[7] There is, however, a line of analysis that takes *-ever* FRs to be universal quantifiers (Bresnan & Grimshaw 1978, Iatridou & Varlokosta 1998). Under this analysis, *-ever* FRs are irrelevant for discriminability because they are quantificational and not entity denoting. We come back to this analysis in the next chapter.

Going back to the examples in (30-32), we see that *–ever* FRs cannot be compared with the expressions discussed in section 1, because if either a proper name or a sortal concept are available, an *-ever* FR cannot be used; knowing one of these counts as "knowing the identity of the entity". Note that Dayal's (1997) representation of "not knowing the identity" of the referent, which is adopted by both von Fintel (2000) and Tredinnick (2005), only deals with the former. In particular, it requires the referent to vary across the worlds compatible with the epistemic alternatives of the speaker in which the FR is evaluated. Authors differ with respect to which are the relevant possible worlds, but the lack of identity across worlds is represented as in (33).

(33) $\iota x[P(w')(x)] \neq \iota x[P(w'')(x)]$

This requirement means that the FR must denote a different entity in each world, i.e. it cannot be rigid. This restriction is fulfilled as long as the speaker does not know the proper name of an entity. Crucially, it will be fulfilled even if the speaker knows the sortal concept of the entity. That is, these analyses predict, counterfactually, that an *-ever* FR can be used in a context in which the speaker knows the sortal concept of the entity. To capture these cases as well, one would have to add a further restriction. Specifically, it should make sure that across the epistemic alternatives of the speaker the entity varies in its sortal concept (or subtypes of sortal concepts – see below). Formalizing this aspect of "not knowing the identity" is left for future research.

Going back to the choice of referring expressions, it should be noted that some *-ever* FRs are possible even if the speaker knows the sortal concept. In particular, it is those -

ever FRs that have heads[8]. For example, in (34) the speaker knows that the person who brought the lasagna is a woman, but is nonetheless ignorant as to her identity. In (35) the speaker knows that Dan went to a resort, but does not know which one. In (36) the speaker knows that the bowl contains soup, but does not know which kind.

(34) Q Who brought the lasagna?
A Whatever woman came during the storm.

(35) Q Where did Dan go for Christmas?
A To whatever resort he went last year.

(36) Q What's worth eating?
A Whatever soup is in the blue bowl.

That is, when a head noun occurs, the indeterminacy persists, but now it concerns a subtype of that lexical noun[9] (although, like with *-ever* FRs with no heads, it might concern the identity of the specific entity). Can these be compared with definite descriptions? It seems like the two are used in the same contexts. However, note that the two types of expressions create different implications regarding what is not known: a

---

[8] Not all *wh*-words allow for head nouns in -ever FRs. In particular, these are only possible with *whatever*, *whichever*, and *however many*. Note, however, that *whoever* and *wherever* FRs can be replaced with *whatever* or *whichever* FRs with an appropriate head noun.

(i) a. I want to meet [whoever (*person/girl/student) gave you the flowers].
b. I want to meet [whatever/whichever person/girl/student gave you the flowers].
(ii) a. I will have lunch [wherever (*place/city / restaurant) she has lunch].
b. I will have lunch in [whatever/whichever place/city/restaurant she has lunch].
(iii) I will read [whatever (book(s)) John reads].
(iv) I will read [whichever *(book) John is reading].
(v) a. I want to weigh [however much Sue weighs].
b. I want to weigh [however many kilos Sue weighs].

[9] Interestingly, given an elaborate enough context, the subtype can be totally ad-hoc. Consider, for example, the following *-ever* FR.

(i) I'll sit on whatever chair you're sitting on.

Out of the blue, this means either that I'll sit on the same chair or on the same kind of chair. But we can construct a context in which you number the chairs in your house, say 1 through 6, and I number the chairs in my house, say 1 through 8. To feel closer, every time we talk on the phone I'll sit on whatever chair you're sitting on, namely, if you're sitting on number 4, I'll sit on number 4 as well. How to capture this kind of property in the meaning of *-ever* FRs is left for future research.

definite description can be used when the speaker knows the identity of the referent but assumes that the addressee does not, whereas the *-ever* FR can only be used if the speaker herself does not know the identity of the referent – the knowledge state of the addressee is not relevant. This is supported by the contrast between the description in (37a) and the *-ever* FR in (37b).

(37) a. I went to the city I grew up in.
b. #I went to whatever city I grew up in.

Since the property expressed in the relative clause makes it necessary that the speaker would know the identity of this place, using an *-ever* FR is awkward at best, while the corresponding definite is fully acceptable, since it implies that the addressee is assumed not to know where this place is.

This distinction is also relevant for comparing standard FRs with *-ever* FRs that lack heads. While both encode the same information about the entity, they imply a different state of the knowledge state of the interlocutors.

(38) a. I went to where I grew up.
b. #I went to wherever I grew up.

To conclude, it is part of the core meaning of *-ever* FRs the speaker does not know the identity of the entity; with standard FRs this is just an implicature, and hence can be violated in favor of a different message. For example, in the location example the speaker can use a standard *where* FR if she wishes to emphasize that Dan always goes to the same place – an *-ever* FR cannot be used in such a situation.

This means that *-ever* FRs cannot be compared with other description because they are not licensed in the same contexts. Independent of whether we identify the indeterminacy with attributivity or not, we see that *-ever* FRs are not simple referential terms and hence are irrelevant for the pattern considered in this chapter. In the next

chapter we discuss *-ever* FRs in the context of specificational sentences.

4. Conclusions

This chapter has introduced new data dealing with the choice of referring expressions and has analyzed it as a scalar phenomenon. We saw that if the common ground contains information that allows constructing more than one referring expressions for a certain entity, the referring expression that must be used is the one that best reflects the familiarity of the interlocutors with that entity. This is in accordance with Grice's Maxim of Quantity, which requires the cooperative speaker to provide all the information available to her; here this information concerns the familiarity of the interlocutors with the entity discussed, which translates into how well the entity is identified, i.e. how well it can be discriminated from other entities. Failing to do so will be interpreted as the speaker trying to convey an extra message in the form of a flouting implicature.

I have proposed that this phenomenon is one of the "clues" that, as proposed by Stalnaker (1978), are dropped during conversation about what each interlocutor assumes about the state of the common ground – it is an indirect way for interlocutors to keep the context non-defective. As such, the choice of referring expressions can be used to study more closely when and how interlocutors coordinate their common grounds.

This phenomenon is also interesting because it is very different from standard scalar implicature which concerns the information asserted. Here, by contrast, the implicature concerns the state of the common ground. This is reminiscent of Gundel et al.'s (1993) analysis of the choice of anaphoric expressions based on their cognitive statuses.

In the next chapter we return to the topic of specificational sentences and use the notion of discriminability to study them.

Chapter 5
Discriminability in Copular Sentences

The previous chapter introduced the notion of discriminability. We saw that while different expressions may be equally successful at establishing reference, some expressions are nonetheless preferred to others. I have argued that the preferred expression reflects a higher level of familiarity of the interlocutors with the referent. We called expressions that reflect more familiarity 'more discriminate' because the more familiar interlocutors are with a certain entity, the better they can discriminate it from other entities. I have proposed that proper names are more discriminate than headed descriptions with contentful nouns, which are more discriminate than headed descriptions with bleached nouns, which, in turn, are more discriminate than (standard) FRs. While I have not provided a definition of discriminability, the idea is that this ranking is a result of the information encoded in the expressions. In particular, knowing the proper name of an entity counts as being more familiar with that entity than knowing its sortal concept which, in turn, counts as being more familiar with the entity than just having number and/or gender information about that entity.

This chapter investigates the notion of discriminability in the context of copular sentences, especially specificational sentences. Our starting point is that specificational sentences are identity sentences (see again chapter 3). The goal is to use discriminability to shed light on what makes certain identity sentences specificational. I will propose that the post-copular phrase in a specificational sentence is the more discriminate expression in the identity. In examining specificational sentences, we will compare them to predicational sentences – especially to predicational sentences which express identity as a

special case. Applying discriminability to copular sentences raises new questions regarding the nature of this notion. While we will see that the notion of discriminability should be relativized both to the specific discourse and to the nature of the entity itself, a complete analysis of discriminability in copular sentence will not be possible. Nonetheless, we will examine the implications of this route to the typology of copular sentences and show the value of using discriminability in analyses of the phenomenon.

The chapter is organized as follows. Section 1 discusses specificational and predicational sentences in the context of familiarity with entities and the discriminability of expressions. I argue that specificational sentences have rising discriminability whereas those predicational sentences that express identity have decreasing discriminability. Since the proposed analysis of discriminability takes predicational and specificational sentences to be opposite in some sense, in section 2 we compare it with the inverse predication analysis of specificational sentences. I argue that specificational sentences express identity and not inverse predication. In section 3 we go back to discriminability and discuss the implications of this notion to the typology of copular sentences.

1. The Discriminability Analysis

In chapter 3 we saw that Higgins claimed that the pseudocleft in (1) is unambiguously specificational because the pre-copular FR is inherently non-referential, i.e. it cannot pick out an entity. The pre-copular position of a predicational sentence requires a referential expression. Consider (1), repeated from chapter 3 (example 20).

(1) What I don't like about John is his tie.

We can demonstrate that this sentence is indeed unambiguously specificational by applying three of Higgins' tests for distinguishing predicational and specificational

sentences: (2a) shows that it is impossible to coordinate an adjectival predicate with the post-copular phrase, (2b) shows that the post-copular phrase cannot be deleted in a coordination of two sentences and (2c) shows that in such a coordination the copula cannot be deleted either. All these are characteristics of specificational sentences – see again chapter 3, section 2.1.

(2) a. *What I don't like about John is his tie and is very expensive.
b. *What I don't like about John is his tie, and what I don't like about Bill is too.
c. *What I don't like about John is his tie, and what I don't like about Bill … his tie.

While Higgins attributes the lack of the predicational reading to certain properties of the FR, it is possible to create a referent for this FR in the context of a guessing game where one has to guess what it is that I don't like about John. In this context, I can utter the unambiguously predicational (3) as a hint.

(3) What I don't like about John is dotted.

The grammaticality of (3) shows that the reason (1) lacks a predicational reading is not that the FR cannot occur in the pre-copular position of a predicational sentence. So maybe the explanation lies in the nature of the post-copular phrase: it could be that the post-copular phrase in (1) cannot be a predicate. However, (4) is a example for a predicational sentence in which *his tie* serves as the predicate.

(4) This piece of fabric is his tie[1].

Again, we can demonstrate that this sentence is indeed predicational by applying Higgins' tests: we see that we can coordinate *his tie* with an adjectival predicate (5a) and

[1] Note that this sentence is only appropriate when the entity referred to by the pre-copular phrase is not easily identifiable as a tie, e.g. if it is not a standard tie but rather some unusual piece of fabric that John uses as a tie. We come back to this in section 1.4.2.

delete both the post-copular phrase and the copula in an appropriate coordination (5b,c). These are characteristics of predicational sentences.

(5) a. This piece of fabric is his tie and is expensive.
b. This piece of fabric is his tie and that piece of fabric is too.
c. This piece of fabric is his tie and that piece of fabric … his handkerchief.

If the expression in the pre-copular position of (1) can occur in the pre-copular position of a predicational sentence (as in 3) and the expression in the post-copular position of (1) can occur in the post-copular position of a predicational sentence (as in 4), why isn't (1) predicational? A potential explanation might be that there is no context for which the predicational meaning is appropriate. However, it seems appropriate in the context of the guessing game that after presenting hints in the form of (3), one could use the predicational reading of (1) to reveal the secret: the sentence would assert that the entity denoted by the FR *what I don't like about John* has the property of being his tie. Nonetheless, (1) does not seem to have this reading.

I propose that the reason (1) is unambiguously specificational is the relation of these two phrases around the copula. More specifically, the relative discriminability of the phrases around the copula, i.e. the level of familiarity with the entity they indicate, is appropriate for specification but not for predication. In the previous chapter we saw that headed descriptions reflect more familiarity of the interlocutors with the entity than FRs, i.e. these are generally more discriminate expressions. I therefore propose that what makes specificational sentences specificational is rising discriminability. That is, specificational sentences are identity sentences in which a certain pragmatic relation holds between the phrases around the copula: the post-copular expression is more discriminate than the pre-copular expression. By contrast, the unavailability of the predicational reading for (1) suggests that if a predicational sentence expresses identity,

discriminability will be decreasing, i.e. the opposite pragmatic relation will hold. Since (1) has rising discriminability, it is specificational and cannot express identity as a special case of predication.

The rest of section 1 develops this proposal. But, first, we take a step back and clarify the assumptions about the semantics of these two types of copular sentences and how these relate to the discriminability analysis.

1.1. *Specificational and predicational sentences*

We saw in chapter 3 that the direct compositional approach to connectivity requires that specificational sentences express identity at the sentence level. But, as we saw in chapter 1, identity can be composed in a number of different ways. Heycock & Kroch (1997, 1999) argue that English specificational sentences are identity statements, and I have argued the same for pronZ pseudoclefts in Hebrew (Heller 1999, chapter 3). Heycock & Kroch derive identity in specificational sentences in English using identity small clauses. For Hebrew, I have argued that pronZ is a '*be* of identity'. Since what is important for my purposes here is that specificational sentences are identity sentences, either analysis would be suitable, and I will not discuss these different proposals here. For the sake of explicitness, I will assume that English also has two *be*'s: '*be* of predication' (λP.P at type <<e,t>,<e,t>>) and 'be of identity' (λYλX.X=Y at type <X,<X,t>>, where X can be of any type).

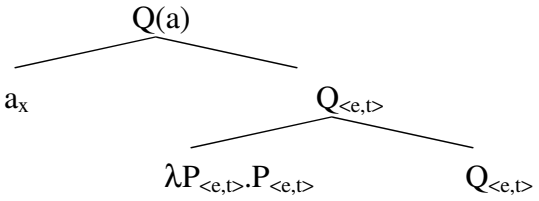
Now, let us compare discriminability in identity sentences to what we saw in the previous chapter with respect to the choice of referring expressions. In both cases, relative discriminability of expressions is relevant because the expressions denote one and the same entity. When choosing a referring expression like we did in the previous

chapter, the common ground contains information that allows for the construction of more than one referring expression for the same entity. In an identity sentence, by contrast, the fact that two kinds of information are associated with the same entity is asserted in the sentence, i.e. this information is added to the common ground. What is available in the common ground prior to that is – to be colloquial – the two entities with their expressions, i.e. each of the two entities is associated with information that allows for the construction of one of the expressions. Note that if this information is not already in the common ground, it will be accommodated – it is presented as given and must be non-controversial.

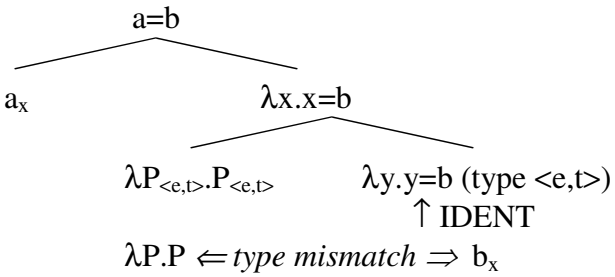
The new claim regarding specificational sentences then is that the pre-copular expression has to be less discriminate than the post-copular expression, i.e. the sentence has rising discriminability. We can also think of specificational sentences as identifying a less familiar entity with a more familiar entity.

Turning to predicational sentences, it is standardly assumed that the pre-copular phrase picks out an entity and the post-copular phrase is a property that is predicated of that entity. Following Partee (1986), I assume that the copula in these sentences does not contribute information – it only guarantees that the post-copular phrase is indeed a property: it is the identity function on properties $\lambda P.P$ (P is a variable at type <e,t>). The semantic composition of simple predication is illustrated in (6a). If the post-copular phrase does not denote a property, its meaning will be shifted into a property if possible. For example, if the post-copular phrase denotes an entity, it can be shifted into denoting a singleton predicate using Partee's (1987) IDENT type shifting rule. (6b) illustrates the semantic composition in this case.

(6) a. Predication



b. Identity as a special case of predication



Thus, predicational sentences can express pure predication, or identity as a special case of predication.

How does discriminability fit into the picture? When the post-copular phrase is a true predicate, discriminability is irrelevant because the sentence does not contain two expressions that denote the same entity. When the sentence expresses identity, we can compare the two phrases for discriminability like in other identity sentences. In these cases, the pre-copular phrase must be more discriminate. That is, the entity must be picked out using the better discrimination available and is said to have the property that it is identical to the entity which is picked out using a less discriminate expression. Decreasing discriminability then does not define predicational sentences in the same way that rising discriminability defines specificational sentences – it is only relevant for a subset of these sentences.

The proposal is summarized in Table 5.1.

Sentence Type	**Semantic Relation**	**Discriminability (linear order)**
Predicational	1. predication	n/a
	2. identity X=Y (as a special case)	X > Y
Specificational	Identity X=Y	X < Y

Table 5.1: Proposed Discriminability in Copular Sentences

It is worth noting that while discriminability in predicational sentences seems to be the opposite of what we find in specificational sentences, these sentences types are not the reverse of each other. This is due to the fact that they do not express an opposite semantic relation, which, in turn, means that discriminability is relevant only for a subset of predicational sentences.

1.2. *Free relatives and headed descriptions*

1.2.1. Post-copular definite descriptions

While Higgins has argued that (1) is unambiguously specificational because of the nature of the FR, I have proposed that this is actually due to the relative discriminability of the phrases around the copula. To test this proposal, we turn to a copular sentence which, unlike (1), has a FR that can clearly refer. The relative discriminability of the phrases around the copula in (7) is the same: we saw in the previous chapter that FRs are generally less discriminate than headed descriptions.

(7) What I'm thinking about is John's watch.

Let us go back to the context of a guessing game, where I'm thinking about some object and you have to guess what it is. In this context I can give clues using predicational sentences similar to (8).

(8) What I'm thinking about is made of plastic.

If I now wish to reveal my secret, I can utter (7). On the predicational reading, the entity denoted by the pre-copular FR should be attributed the property of being John's watch. On the specificational reading, this entity is equated with the entity denoted by the post-copular description.

We can test which readings are available in (7) by applying Higgins' tests, as in (9). (9a) shows that a predicate cannot be coordinated with the post-copular phrase, (9b) shows that the post-copular phrase cannot be deleted in a coordination, and (9c) shows that the copula cannot be deleted in a similar coordination of two sentences.

(9) a. *What I'm thinking about is John's watch and is his favorite accessory.
b. ??What I'm thinking about is John's watch and what Rachel is thinking about is too.
c. *What I'm thinking about is John's watch and what Rachel is thinking about … John's tie.

That is, abstracting from Higgins' original FR, we still find the same pattern: (7) is unambiguously specificational. This example supports my proposal that if an identity sentence has rising discriminability it will be specificational and cannot be predicational.

The same pattern is found in Hebrew, which – as we saw in chapter 3 – distinguishes predicational and specificational sentences in the choice of the copula. Recall that a copular sentence with a pronH copula is predicational and a copular sentence with a pronZ copula is specificational. The Hebrew example in (10) is the Hebrew equivalent of (7). As expected, only pronZ is available, which means that this sentence is unambiguously specificational.

(10) ma še-ani xoševet alav *hu/ze ha-sha'on šel dan
what that-I think about-him H/Z the-clock of Dan
'What I'm thinking about is Dan's clock.'

As with (1), we can show that the reason for the ungrammaticality of the pronH version is that the two phrases around the copula cannot occur in a pronH sentence <u>together</u> – each

of the phrases can occur in the same position in a different pronH sentence. In particular, (11a) shows that the FR in (10) can occur in the pre-copular position in a pronH, i.e. predicational, sentence and (11b) shows that the post-copular phrase in (10) can occur in a post-copular position in a predicational sentence.

(11) a. ma še-ani xoševet alav hu mataxti
what that-I think about-him H metallic
'What I'm thinking about is metallic.'

b. xatixat ha-matexet ha-zot hi ha-sha'on šel dan[2]
piece the-metal the-this H the-clock of Dan
'This piece of metal is Dan's watch.'

The Hebrew pattern is similar to what we saw above for English: the pronZ version is grammatical because the relative discriminability in the sentence is rising. Such rising discriminability is incompatible with a predicational sentence and hence the pronH version of (10) is ungrammatical.

1.2.2. Post-copular indefinite descriptions

When discussing the choice of referring expressions in the previous chapter, we did not consider indefinite descriptions because when an indefinite is the appropriate referring expression, the entity will either not be in the common ground (Kamp 1981/Heim 1982) or will not be associated with other information such that other referring expressions can be constructed. Nonetheless, I have proposed that the information encoded in indefinites should make them as discriminate as definites. They also express the sortal concept of

[2] Like in (4), the entity in (11b) is such that it is not easily recognizable as a watch – maybe it has a weird shape. We come back to this in section 1.4.2.

the entity as well as its number and gender. This section examines discriminability in specificational sentences with indefinites.

We begin by considering the indefinite version of (10). In the context of the guessing game in which I'm thinking about some clock and (12) is intended to reveal my secret, only pronZ is possible. That is, the sentence can be specificational but not predicational.

(12) ma še-ani xoševet alav ??hu/ze sha'on
what that-I think about-him H/Z clock
'What I'm thinking about is a clock.'

The availability of the specificational version but not the predicational one is the pattern we saw for the definite version in (10) – it suggests that indefinites are indeed of the same discriminability as definites.

Interestingly, the pronH version in (12) is grammatical in a different context. In particular, if the entity I am thinking about is a microwave, I can utter (12) as a hint. In this context, the entity picked out by the pre-copular FR is not a clock but rather has a clock as a property of it. In this context, (12) is not an identity sentence, because the post-copular indefinite *sha'on* 'a clock' is not an alternative label for the entity denoted by the pre-copular FR. Rather, the pronH version of (12) is a standard predicational sentence and does not express identity at any level. Although the predicate is nominal which can denote an entity in other contexts, here it is a predicate similar to the adjective in (11a). Crucially, when referring to such an entity, relative discriminability is not relevant. Note that in this context the specificational version is simply false – what I'm thinking about is not identical to a clock.

The example in (12) behaves as predicted if indefinites are of the same discriminability as definites. But this is not the whole picture. In particular, in chapter 3 we saw an example of a pseudocleft which, like (12), has rising discriminability and yet it

allows for both copulas, i.e. it was said to be ambiguous between predicational and specificational. Consider (13), repeated from chapter 3 (example 11).

(13) ma še-ani macbi'a alav hu/ze xatul
what that-I point on-it H/Z cat
'What I'm pointing at is a cat.'

Let us first remind ourselves of the intuitive meaning of the two versions. On the specificational reading, I am telling the addressee in which direction I'm pointing – there is some cat which is where I'm pointing. On the predicational reading, on the other hand, the FR picks out the "the nearest object intersected by the line formed by producing the longitudinal axis of my forefinger" (Higgins, p.212) and the sentence says of that object that it is a cat.

Prima facie, the availability of the predicational reading in (13) constitutes a counter-example for the discriminability analysis: the sentence has rising discriminability (FR << description) which I propose to be only compatible with specificational sentences. It also seems that this case is not similar to the alternative context for (12) (the microwave situation) in which the pronH versions was possible. In particular, in that context the post-copular phrase in (12) did not provide an referring expression for the entity denoted by the pre-copular phrase, while here it does. Nonetheless, I propose that this case is similar in that the post-copular phrase is a property and not an alternative description. This is because the predicational reading here is only felicitous if the addressee cannot identify the entity in question as a cat, either because (i) the specific instance looks very different from the standard cat, or (ii) the addressee doesn't know at all what a cat is. Since the sortal concept is new, it is not an alternative description of the entity and hence discriminability is irrelevant. That is, a novel sortal concept functions like a run-of-the-mill predicate. We will see in section 3.1 that sentences in which the post-copular

predicate is an inherent property are identificational. That is, (13) is actually both identificational and specificational and not predicational and specificational as originally proposed by Higgins.

Why is such a reading possible in (13) and not in (12)? The difference lies in the nature of the FR. The FR in (13) picks out the entity by ostension: its physical presence makes it appropriate to teaching a new concept, while in the other cases the entity is not made physically available and hence its properties are not available such that the addressee will learn a new sortal concept. This is not special to Hebrew: the original example was given by Higgins in English – consider the ambiguous (14).

(14) What I'm pointing at is a cat.

Like (13), (14) is ambiguous between the same predicational and specificational readings.

The contrast between (12) and (13-14) shows that nominal predicates fall into two categories: nominals that encode the sortal concept of the entity (which we colloquially refer to as "what the entity is") and nominals that express non-essential properties of the entity and hence can be viewed as "true predicates". While the former may be compared with other expressions for relative discriminability, the latter may not because they are not potential referring expressions for the entity. Note, importantly, that this distinction was not found with definites. This is because in a definite the head noun always encodes the sortal concept information so it cannot be a "pure predicate". In addition, definites are not appropriate for introducing new sortal concepts – using a definite presupposes identifying the entity in question as the unique N which requires identifying it as an N.

1.2.3. New vs. mutual information: an alternative?

The different readings we saw for indefinites in the previous section might suggest that the availability of predicational and specificational readings is not related to discriminability. Instead, one may suggest that the possible readings of copular sentences relates to a distinction between new vs. given information, i.e. what is added to the common ground vs. what is already given. In particular, mutual information will yield a specificational reading and new information will yield a predicational reading.

This generalization, however, only captures the last two cases we discussed in which a post-copular indefinite introduces a new sortal concept. Recall that for (1), (7) and (10) the predicational reading was not available independent of whether the definite in the post-copular expression expresses is old information, e.g. whether it is already known that John has a tie (for 1). In (12) where the post-copular phrase is indefinite, it is only assumed that the interlocutors know what a clock is, not whether some clocks are available in the context.

In sum, we see that if certain kinds of information are available about an entity, "more specific" information is added in a specificational sentence and "less specific" information can be predicated. This is reflected in the relative discriminability of expressions: a more discriminate expression contains information that reflects more familiarity with the entity and a less discriminate expressions contains information that reflects less familiarity with the entity. Note that this analysis explains why FRs are commonly found as specificational subjects: since FRs are relatively low in their discriminability, they will occur in a position that requires lower discriminability.

1.3. *Descriptions and proper names*

When looking at referential uses of proper names in the previous chapter, we saw that proper names are the most discriminate expressions. If we straightforwardly apply this to copular sentences, we expect (15a) to be unambiguously predicational and (15b) to be unambiguously specificational.

(15) a. Dan is my next-door neighbor.
b. My next-door neighbor is Dan.

In order to determine which readings are available for these sentences, we again use Higgins' tests for distinguishing predicational and specificational sentences. The (a) examples test whether a predicate can be coordinated, the (b) examples test whether the post-copular phrase can be deleted in a coordination and the (c) examples whether the copula can be deleted in a coordination.

(16) a. Dan is my next-door neighbor and is Italian.
b. Dan is my next-door neighbor and Bill is too.
c. Dan is my next-door neighbor and Bill … my hairdresser.

(17) a. *My next-door neighbor is Dan and is Italian.
b. *My next-door neighbor is Dan and my hairdresser is too.
c. *My next-door neighbor is Dan and my hairdresser … Bill.

The tests in (17) indeed indicate that (15b) is unambiguously specificational (but see 20 below). The tests in (16), on the other hand, only show that (15a) has a predicational reading. Nothing can be concluded regarding a specificational reading for it. This is because the availability of the predicational reading is sufficient to allow for the sentence to "pass" the tests and this may "hide" the ungrammaticality that may be caused by the specificational reading. Unfortunately, all of the tests that examine the sentence as a whole give rise to the same pattern. Note, importantly, that tests which target only one element are not sufficient – we have already seen that it is the combination of elements

that makes a sentence predicational or specificational. As for (15a), we cannot reach a conclusion regarding whether it has a specificational reading. But for (15b), the tests show the discriminability analysis is correct: the sentence has rising discriminability (description << proper name) and hence is specificational.

Turning now to Hebrew, we first examine the pronH and pronZ versions that correspond to (15b). Given what we saw for English, we expect the pronZ version, i.e. the specificational sentence, but not the pronH version, i.e. the predicational sentence, to be grammatical. The actual pattern is more complex: while the pronH version is indeed odd at best with a first name alone, it becomes acceptable with a full name.

(18) a. ha-more šeli le-karate hu dan ??(kohen)
the-teacher mine to-karate H Dan Cohen
'My karate teacher is Dan (Cohen).'

b. ha-more šeli le-karate ze dan (kohen)
the-teacher mine to-karate Z Dan Cohen
'My karate teacher is Dan (Cohen).'

I propose that the acceptability of the full name version in (18a) is reminiscent of the introduction of a new sortal concept with an indefinite. In particular, when a name is predicated over an entity, it amounts to introducing this name into the common ground – since it is not yet available in the common ground, it cannot be compared for discriminability with the description in the pre-copular position. This is supported by the contrast with the first name, which is used when the name is already familiar, i.e. when the proper name is already available in the common ground.

In fact, the same pattern is found in English. If we add a last name to (15b), as in (19), the sentence passes all three tests – this is demonstrated in (20).

(19) My next-door neighbor is John Smith.

(20) a. My next-door neighbor is John Smith and is British.
b. My next-door neighbor is John Smith and my hairdresser is too.
c. My next-door neighbor is John Smith and my hairdresser – Bill White.

The fact that the proper names here function as predicates is especially clear when examining the meaning of (20b) – it means that my next-door neighbor and my hairdresser happen to have the same name, not that they are in fact the same person. With a first name only in (17), this reading was not available.

Next, we turn to the Hebrew counterparts of (15a). As we see in (21), both the pronH version, i.e. the predicational sentence, and the pronZ version, i.e. the specificational sentence, are grammatical. These two sentences are truth-conditionally equivalent: they both means that the individual denoted by the proper name *Dan* and the individual denoted by the definite description *ha–more šeli le-karate* 'my karate teacher' are one and the same. Nonetheless, the two sentences differ in their non truth-conditional meaning.

(21) a. dan hu ha-more šeli le-karate
Dan H the-teacher mine to-karate
'Dan is my karate teacher.'

b. dan ze ha-more šeli le-karate
Dan Z the-teacher mine to-karate
'Dan is my karate teacher.'

The intuitive meaning difference is that in the predicational sentence in (21a) the pre-copular phrase is mutual knowledge, while in the specificational sentence it is the post-copular phrase that is the mutual knowledge. In (21a), the interlocutors already have an established referent for the name *Dan*, and by predicating the post-copular expression over this individual, the speaker adds another potential expression that could pick out the same entity. In the specificational (21b), the information that the speaker has a karate teacher is presented as known (or else it can be accommodated). The assertion made by

the speaker is that the proper name *Dan* has a more familiar instantiation in the common ground as the karate teacher.

I propose that this meaning difference follows from the relative discriminability of the expressions that is dictated by the sentence type, i.e. by the copula. In particular, the specificational (21b) forces the proper name to be less discriminate than the description – this is the opposite of what we saw for a situation when both were available in the common ground. (21b) will be appropriate as a clarification in a context where the proper name has been used and has turned out not be known to the addressee – in this context, (19b) identifies the less familiar referent of the proper name with the more familiar referent of the description.

The flexibility of discriminability with proper names as opposed to what we saw with descriptions is due to the fact that the association of proper names with their referents is arbitrary, unlike the association of descriptions which depends on the entity having the property mentioned in the description. This difference affects the way proper names can be introduced. In particular, introducing new proper names in this way, i.e. in (18a) and (21b), is much more common than introducing new concepts – while names are just tags that identify a certain individual, sortal concepts are complex so they are not normally introduced in this way.

A refinement in the notion of discriminability is called for. It is not a fixed notion; we cannot say that proper names are by definition more discriminate than headed description (this was always the case in the previous chapter because contexts were controlled to ensure that the expressions were all appropriate ways of referring to the entity, given the common ground. The relative discriminability of expressions depends on the familiarity

of the interlocutors with entities in the specific context in which discriminability is evaluated.

1.4. *Two headed description*

1.4.1. Bleached nouns

In the previous chapter we have drawn a distinction between headed descriptions with "bleached" head nouns and those with lexically contentful head nouns based on the differences found between them in the choice of referring expressions. This section examines whether the same distinction can be carried over to specificational sentences.

Consider (22) in which the pre-copular phrase is headed by a bleached head noun *thing* and the post-copular phrase is headed by a lexical noun. According to what we saw in the previous chapter, we expect this sentence to have rising discriminability and hence be unambiguously specificational.

(22) The thing I want to eat is the soufflé.

Again, we apply Higgins' tests to examine which sentence types are available here. The tests in (23) confirm that the sentence in (22) is unambiguously specificational, as expected: a verbal predicate cannot be coordinated (23a) and neither the post-copular phrase nor the copula can be deleted in an appropriate context (23b,c).

(23) a. *The thing I want to eat is the soufflé and should be delicious.
b. *The thing I want to eat is the soufflé and the thing Dan wants to eat is too.
c. *The thing I want to eat is the soufflé and the thing Dan wants to eat - the mousse.

However, other copular sentences in which the pre-copular phrase is headed by a bleached noun and the post-copular phrase is headed by a lexical noun do not yield such a clear pattern. Consider the copular sentences in (24) and (26) in which the pre-copular phrase has a bleached head and the post-copular phrase has a contentful head – the tests

are given in (25) and (27) respectively.

(24) The guy with the red shirt is the next-door neighbor.

(25) a. The guy with the red shirt is the next door neighbor and makes excellent lasagna.
b. ?The guy with the red shirt is the next door neighbor and the guy with the pink shirt is too.
c. ?The guy with the red shirt is the next door neighbor and the guy with the pink shirt … the upstairs neighbor.

(26) The place where I live is the capital of England

(27) a. The place where I live is the capital of England and is beautiful
b. ?The place where I live is the capital of England and the place where you live is too.
c. The place where I live is the capital of England and the place where you live … the capital of France.

Both (24) and (26) allow coordinating a predicate (25a, 27a) which is a characteristic of predicational sentences. Deleting the post-copular phrase yields a marginal result for both sentences – note that in the case of (25b) this may be due to the fact that two different guys may not be the same next door neighbor. Finally, deleting the copula is marginal for (25c) and perfectly possible for (27c), again suggesting that these are predicational. In sum, the result of these tests are not conclusive, but these sentences are clearly not specificational as predicted.

This could indicate that the distinction made between bleached and contentful nouns in discussing the choice of referring expressions was not on the right track. Alternatively, in discussing the choice of referring expressions we abstracted from the role of modification, but it is likely that it does play a role in determining the discriminability of an expression. Since our understanding of discriminability at this point is limited, I do not know how to factor modification in. I have included this here for completeness and defer further investigation for later.

1.4.2. Nouns with no logical relation

There are also specificational sentence with two headed descriptions in which the head nouns do not stand in any logical relation to each other. The example in (28), due to Higgins (1973, p. 245)), is ambiguous between a predicational and a specificational reading.

(28) Nixon's plan is a bomb.

The specificational reading talks about a bomb which constitutes Nixon's plan. The predicational reading says that Nixon's plan is dangerous. The indefinite *a bomb* is a property of the entity denoted by the pre-copular expression and not an alternative description of it. A second example, also due to Higgins, is given in (29).

(29) His hat is a bundle of straw.

The specificational reading of (29) means that he uses a bundle of straw as a hat – in some sense, it is not "really" a hat. On the predicational reading, the sentence talks about a "real" hat which is made of bundled straw.

The predicational readings of these sentences further show that the not all nominal predicates encode the sortal concept of the entity denoted. Thus, these readings are irrelevant for discriminability – this is similar to the clock example in (12). Note that this usage of is not limited to indefinite descriptions. In particular, we saw earlier in this chapter examples in which the post-copular predicate was definite. (30a) is repeated from (4) and (30b) is repeated from (11b).

(30) a. This piece of fabric is his tie.

b. xatixat ha-matexet ha-zot hi ha-sha'on šel dan
piece the-metal the-this H the-clock of Dan
'This piece of metal is Dan's clock.'

I have noted that the referent of the subject in (30a) is not a standard tie – it is only used as a tie, i.e. tie is a property of the referent and not an alternative description of it. The same is true of the referent of the subject in (30b): it is "really" a piece of metal that is used as a clock, i.e. clock is a property of it, not its essence. This accounts for why these expressions serve as predicates here while they could not serve as predicates for other entities in sentences like (1) and (12).

Going back to the specificational readings of (28-29), these sentences show that even the sortal concept an entity is associated with is not unique – the same entity can be conceptualized as more than one thing. In (28), the same entity is conceptualized as both a bomb and a plan. In (29), the entity is conceptualized as both a hat and a bundle of straw. If we apply the discriminability analysis to these sentence, it means that in (28) *a bomb* would be more discriminate than *Nixon's plan* and in (29) *a bundle of straw* would be more discriminate than *his hat*. While it is not straightforward how this can be formalized, there is an intuitive feeling that under the specificational reading of (28) the entity in question is somehow "more a bomb than a plan" in the sense that it is as a bomb that it constitutes the plan. The same intuition is found with (29): the entity in question is "more a bundle of straw than a hat" because it is a "real" bundle of straw which only serves as a hat.

These examples show that nominals can have different roles. First, there are nouns which are predicates and not sortal concepts. And not all sortal concepts are equal – some are in some sense more inherent than others, even though they all can be used as heads in referring expressions for the entity (as opposed to true properties). These distinctions are relevant for discriminability. In particular, they show that discriminability has to be

relativized to the specific entity in question: the same expression may be more discriminate when denoting one entity and not the other.

Note, interestingly, that distinguishing between sortal and non-sortal information and also among different sortal concepts is not specific to specificational sentences. This distinction is relevant for representing nominal predicates in general. I hope to take on this issue in future work.

1.5. *–ever free relatives*

While both standard FRs and *-ever* FRs can occur in the pre-copular position of a predicational sentences, as in (31), Jacobson (1995) has noted that *-ever* FRs are impossible in the same position in a specificational sentences, as in (32) (see again chapter 2 section 2.1).

(31) a. What Mary bought was expensive.
b. Whatever Mary bought was expensive.

(32) a. What Mary bought was Barriers.
b. *Whatever Mary bought was Barriers.

One explanation for this pattern was presented by Iatridou & Varlokosta (1998). Their explanation rests on two assumptions: (i) that *-ever* FRs are universally quantified expressions and (ii) that specificational sentences are an instance of inverse predication. If specificational sentences are inverse predication, their pre-copular position is a predicative position in which quantificational expressions, such as *-ever* FRs, cannot occur. While this is an elegant account for the ungrammaticality of (32b), neither of the assumptions it rests on are solid. In section 2 of this chapter I argue against the inverse analysis of specificational sentences. In addition, as mentioned in the previous chapter, Jacobson (1995) has argued in detail that *-ever* FRs are entity-denoting rather than

quantificational. This is also the position of Dayal (1997), von Fintel (2000) and Tredinnick (2005). Since we reject both assumptions made by Iatridou & Varlokosta (1998), we cannot adopt their explanation.

Note, further, that the ban of *-ever* FRs from the pre-copular position of specificational sentence is not total. While an *-ever* FR cannot occur in a specificational sentence if the post-copular phrase specifies the proper name of the entity (as in 32b) the sortal concept of the entity (as in 33a-b) or a sub-sortal concept (as in 33c) it has been pointed out by Dayal (1997) that such a FR can occur in a specificational sentence with a post-copular FR, as in (33d).

(33) a. *Whatever Bill cooks is soup.
b. *Whatever Bill cooks is the soup that his children loves.
c. *Whatever soup Mary is eating is corn chowder.
d. Whatever Bill cooks is what Harry eats.

In order to show that (33d) is indeed specificational, we apply Higgins' structural tests. The tests in (34) confirm that (33d) is specificational.

(34) a. *Whatever Bill cooks is what Harry eats and is always delicious.
b. *Whatever Bill cooks is what Harry eats and whatever Joyce cooks is too.
c. *Whatever Bill cooks is what Harry eats and whatever Joyce cooks … what Beth eats.

Dayal (1997) proposes using the indeterminacy of *-ever* FRs with respect to the identity of the entity to account for the contrast between (32b), (33a-c) and (33d). In particular, if an *-ever* FR means that the speaker "does not know the identity of the entity", the speaker is not expected to have information such as the proper name or the (sub-)sortal concept of the entity, but knowing the information encoded in a FR is allowed[3].

[3] Recall from the previous chapter that the actual formalism proposed by Dayal (1997) (and adopted by von Fintel 2000 and Tredinnick 2005) only captures not knowing the proper name.

Dayal presents interesting evidence for this analysis in the form of a negated specificational pseudocleft with an *-ever* FR contrasting with its positive counterpart. The contrast is attributed to the fact that in the negated pseudocleft the post-copular phrase does not specify the identity of the entity denoted by the *-ever* FR and hence is compatible with the speaker not knowing "the identity of the referent".

(36) a. *Whatever Mary bought was Barriers.
b. ?Whatever Mary bought was not Barriers.

The status of (36b), however, is not uncontroversial. In particular, Jacobson (1995) has considered – and rejected – the pragmatic explanation for the ungrammaticality of (36a) based on the contrast in (37). (37a) differs from (36b) minimally in that the *-ever* FR is clefted – the role of clefting is to emphasize the ignorance reading. Jacobson argues that difference between the ungrammatical (37a) and the grammatical (37b) in which the *-ever* FR is left-dislocated shows that the epistemic state of the speaker is irrelevant: in (37b) the speaker knows that the referent of the *-ever* FR is an apple.

(37) a. *Whatever it was that John ate wasn't an apple.
b. Whatever it was that John ate, it wasn't an apple.

The problem with this contrast is that what seems to be an *-ever* FR in the dislocated position in (37b) is not a FR but rather a clausal adjunct, as argued by Izovrski (2000). Thus, the examples in (37) are actually not a minimal pair. Nonetheless, the different status of (36b) and (37a) is problematic for Dayal's analysis.

One possibility is that the different status of (36b) and (37a) is due to the nature of the post-copular phrase: a proper name in the former and a headed description in the latter. To test this idea, we consider the examples in (38). For (38a), consider a context where you have just seen Bill eating something that resembles an apple, but it is purple so you are not sure. You report this event to me, a good friend of Bill's. Knowing that Bill is

allergic to apples, I can utter (38a). For (38b), you see Mary standing in the kitchen cooking some yellowish soup, and you report this to me as her cooking corn chowder. In this context, I can utter (38b).

(38) a. Whatever Bill is eating is not an apple, because Bill is allergic to apples.
b. Whatever soup Mary is cooking is not corn chowder, because Mary is allergic to corn.

Note that the contexts I have provided for the examples in (38) are such that the speaker does not know the identity of the referent, but has information that guarantees what this referent is not. In this context, my consultants found these sentences totally natural – they did not report even the oddity mentioned by Dayal. It is possible that when presented with no specific context these sentences are odd because it is not a natural situation that the speaker would not know what the referent is and yet know what it isn't.

This creates the following picture for *-ever* FRs in specificational pseudoclefts: (i) they are impossible with a post-copular name or headed description; (ii) they are possible with a FR; (iii) they are possible in a negated pseudocleft with a post-copular description. Based on these facts, I conclude that Dayal's (1997) analysis suggesting that specifying the identity of the referent contradicts the meaning of an *-ever* FR is essentially correct As we saw in the previous chapter, further work is required in order to capture all the cases of "not knowing the identity of the referent" – the exact formalization will affect the predictions of what can and cannot occur in specificational pseudoclefts with *-ever* FRs.

1.6. *Concluding remarks I: discriminability with non-NPs*

Up to this point we have only considered nominal expressions, both when choosing a referring expression and in specificational sentences. But we have already seen in chapter

3 that specificational sentences equate adjectival and verbal expressions in addition to nominal ones. This section presents ideas about how discriminability can deal with non-nominal specificational sentences.

First, in non-specificational environments, we can find different levels of familiarity with properties – this is illustrated in (39-40). If honest is the one thing I want a man to be, the examples in (39) are truth-conditionally equivalent, and if John went home, the examples in (40) are truth-conditionally equivalent.

(39) a. John is honest.
b. John is the one thing I want a man to be.

(40) a. Julia went home.
b. Julia did what John did.

Nonetheless, the intuition is that the adjective in (39a) is in some sense "more specific" than the corresponding "bleached" description of the property in (39b). Similarly, the verb phrase in (40a) is "more specific" than the FR describing the same event in (40b) (I am ignoring the issue of *do*-support). While this intuition is even harder to pin down than what we saw for the case of familiarity with entities, it seems reasonable to say that the AP in (39a) and the VP in (40a) reflect better familiarity of the interlocutors with the property/event. Thus, they are more discriminate expressions, so we expect the same relative discriminability in specificational sentences.

(41) a. The one thing I want a man to be is honest.
b. What John did was go home.

Of course, this is only an illustration that the discriminability analysis can in principle be extended to non NPs. Developing this idea requires a better understanding of discriminability – this is left for future research.

1.7. *Concluding remarks II: the bigger picture*

We started this section with the proposed generalization that specificational sentences are identity sentences in which the post-copular expression is more discriminate than the pre-copular phrase, and that in predicational sentences expressing identity, by contrast, the post-copular phrase must be less discriminate (before undergoing IDENT) than the pre-copular phrase.

The notion of discriminability we started out with was the one introduced in the previous chapter, where different kinds of information had a fixed level of discriminability: a proper name was argued to be more discriminate than sortal information, which in turn was argued to be more discriminate than number and gender information. In examining discriminability in copular sentences we saw that discriminability is not fixed. First, when examining proper names in section 1.3 we saw that discriminability has to be relativized to the specific discourse context – the discriminability of a proper name changes depending on whether or not it was already in the common ground. This is because proper names are arbitrarily associated with their referents – it contrasts with descriptions which require that the entity have certain properties and hence always provide information about it that makes it familiar to some extent.

Second, in examining headed nominals in section 1.4.2 we saw that the level of familiarity expressed by different sortal concepts depends on the nature of the entity – the discriminability of a headed description with a head noun like *hat* depends on whether the entity is a hat, or just being used as a hat (see again example 29). That is, discriminability has to be relativized to the nature of the specific entity with which the expression is

associated.

I therefore conclude that discriminability is a much more fine-grained notion than originally proposed. At this point, I will not provide a definition of discriminability. Nonetheless, this analysis sheds light on the nature of the copular sentences. Table 5.2., repeated in part from table 5.1, summarizes the discriminability analysis and raises questions about other possible combinations. I briefly speculate on one of these below.

Sentence Type	**Semantic Relation**	**Discriminability**
Predicational	1. predication	n/a
	2. identity X=Y (as a special case)	X > Y
Specificational	Identity X=Y	X < Y
???	Identity X=Y	X > Y

Table 5.2: Discriminability in Copular Sentences

The first question is regarding the existence of identity sentences in which the pre-copular phrase is more discriminate than the post-copular phrase, i.e. sentences which exhibit the opposite discriminability pattern to specificational sentences. While we have not discussed these sentences here, it is known that specificational sentences can be reversed, as in (42).

(42) a. What I don't like about John is his tie.
b. His tie is what I don't like about John.

If both discriminability patterns are possible, what is the difference between the two sentence types? Note that reversed specificational sentences have a marked intonation pattern that differs both from that of specificational sentences and from predicational sentences. Characterizing this intonation requires phonological work that is beyond the scope of this dissertation. However, the fact that this intonation seems to be marked may indicate that their discriminability pattern is marked for an identity sentence. This might be explained by the intuition that it is most natural to identify an entity that is less

familiar with an entity that is already better established, i.e. to equate an entity denoted by a less discriminate expression with an entity denoted by a more discriminate expression. When the information in the post-copular phrase is less discriminate, it would be more natural to use this information as a predicate, i.e. to have a predicational sentence, instead of identifying the entity denoted by the pre-copular expression with that entity. An identity sentence with decreasing discriminability is the lest preferred option. This suggests that the relative discriminability in copular sentences is a by-product of constraints on updating information. Discriminability may be better understood if it is studied in the context of a dynamic framework.

2. WHY NOT INVERSION?

The discriminability analysis of specificational sentences proposes that what makes certain sentences specificational is that they are identity sentences with rising discriminability. By contrast, I have proposed that in predicational sentences expressing identity the relative discriminability is the opposite. Taking predicational and specificational sentences to be – in some sense – the opposite of each other reminds us of the inverse analysis for specificational sentences. Given this similarity, it is important to point out that the two approaches are in fact very different. The goal of this section is to compare the inverse predication analysis to the analysis proposed here. I argue that my approach better captures the range of copular sentences.

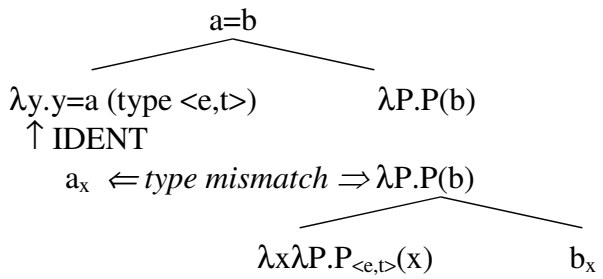
2.1. *Specificational sentences as inverse predication*

The analysis of inverse predication takes specificational sentences to encode the same relation as predicational sentences, i.e. predication, but from a different perspective. The idea is that predicational sentences exhibit the order subject-predicate and specificational

sentences are the opposite, i.e. they exhibit the order predicate-subject. The inverse predication analysis for specificational sentences was first suggested in Williams (1983), and was later adopted in different versions by Partee (1986), Jacobson (1995), Moro (1997) and Mikkelsen (2004b), among others. While analyses differ, they all share two central features: (i) specificational sentences are the opposite of predicational ones, and (ii) the pre-copular position is a predicative position.

The goal of this section is not to provide a review of the inverse analysis, but rather to examine it in the light of the analysis proposed here. So, first, if we wish to account for connectivity using the direct compositional approach argued for in chapter 3, we need to assume that under the inverse predication analysis specificational sentences can express identity at the sentence level. This is rather straightforward: if the grammar contains Partee's (1987) IDENT type-shifting rule the denotation of a pre-copular entity can be shifted to denote a predicate which will be predicated over the over the post-copular entity. Note that we have already assumed the grammar to contain IDENT in order to derive identity as a special case of predication. (43) illustrates how identity is derived as a special case of inverse predication.

(43) *Identity as a special case of inverse predication*



This section presents two central points on which the two approaches differ.

2.2. *Inverting the elements around the copula*

If specificational sentences encode the opposite relation from predicational sentences, we expect that inverting the elements around the copula in a predicational sentence would yield a specificational one. If, on the other hand, specificational sentences are identity statements, only some predicational sentences, i.e. those that can express identity as a special case of predication, should be invertible.

As pointed out by Heycock & Kroch (1999), inversion is not always possible. While (44a) can invert, this is not the case for (44b,c).

(44) a. John is the best candidate for the job. (Heycock & Kroch 1999)
b. John is a doctor.
c. John is proud of his daughters.

(45) a. The best candidate for the job is John.
b. *A doctor is John.
c. *Proud of his daughters is John.

Being aware of these examples, Mikkelsen (2004b) proposes that the ungrammaticality of examples like (45b,c) is due to the discourse function of the inversion that turns predicational sentences into specificational ones. In particular, predicate fronting is only possible if the predicate contains discourse old material. This condition is given in (46).

(46) DISCOURSE CONDITION (STRONG) (p. 225)
The initial element of a DP-inversion must be at least as discourse old as the final element, and it cannot be entirely discourse new.

In (45a), the inverted predicate satisfies the restriction by being a definite. Mikkelsen, who follows work by Prince (1992) and Birner (1996), states that "we can understand the use of definite forms in specificational subjects as an indirect reflex of their Discourse-oldness" (p. 227). That is, if an expression is definite, its referent is discourse old and it can be fronted, i.e. appear in the pre-copular position of a specificational sentence. The expression in the pre-copular position of (45b), in contrast, does not satisfy the discourse

restriction on inversion, since as an indefinite it is discourse new. The prediction for (45c) is not as clear: Mikkelsen does not discuss specificational sentences with a pre-copular AP, but her discussion of predicate fronting of APs suggests that an AP may be Discourse-old if it serves some connecting function (p. 217). It suggests that the ungrammaticality of (45c) may be due to the fact that with no prior context, the AP is not Discourse-old.

Furthermore, Mikkelsen notes that while bare indefinites like (45b) are never possible, there are some indefinites that can occur in the pre-copular position of a specificational sentence. What characterizes these indefinites is that they contain discourse old material in the relative clause and hence can be considered as Discourse-old (Mikkelsen's 8.37).

(47) The occurrence of a factive sentence in contexts like (40) and (41) shows that factives do not always have the force of the-fact-that-φ sentences. ….
A philosopher who seems to share Kiparsky's intuitions on some factive predicates is Unger (1972), who argues that a sentence like (i) entails (ii).

In sum, Mikkelsen's condition is attractive because it seems to address the main concern regarding the inverse approach – namely, why some predicates can invert while others cannot.

There is, however, one class of examples which are not captured by Mikkelsen's constraint, which wrongly predicts them to be acceptable. In particular, if the predicate is a definite predicate (as opposed to a definite which denotes an individual), the sentence cannot be inverted. This is illustrated in (48-49) – the first example is from Heycock & Kroch (1999).

(48) a. John is the one thing I have always wanted a man to be. Namely, honest.
b. *The one thing I have always wanted a man to be is John. Namely, honest.

(49) a. John Smith is what Dan Blum was last year. Namely, the chair.
b. *What Dan Blum was last year was John Smith. Namely, the chair.

One may suggest that inversion is not possible because while the property is indeed definite and hence discourse-old, it might not be "at least as discourse old as the final element". In (50) we thus consider a context for (48) which makes the pre-copular phrase more Discourse-old than the post-copular name. Even in this context, which should support inversion, the grammaticality of the sentences in (48) stays the same. Note, importantly, that this context does support a specificational sentence with the desired pre-copular phrase, as in (50c).

(50) Men are loud and smelly and all kinds of nasty things. But I don't care about all that. There's only one thing I want in a man.

a. John is the one thing I want a man to be. Namely, honest.
b. *The one thing I want a man to be is John. (Namely, honest.)
c. The one thing I want a man to be is honest.

We can thus conclude that the ungrammaticality of (50b) is not due to an impossible inversion of the predicate, but rather due to the relation between the two expressions around the copula. In (50b) the pre-copular phrase denotes a property and the post-copular phrase denotes an individual, whereas in (50c) both phrases denote properties. While these may seem like a single kind of example, they are exactly of the nature for which the two approaches make different predictions. This example thus shows that in a specificational sentence the two phrases around the copula must denote the same kind of entity. We conclude that specificational sentences are identity sentences and not inverse predication.

2.3. *Reversed specificational sentences*

A second difference between the inverse approach and the identity approach is with respect to reversed specificational sentences. If specificational sentences are inverse predication (modulo some constraints), reversed specificational sentences are simply predicational sentences. On the other hand, if specificational sentences are identity sentences with a certain discriminability relation, reversed specificational sentences are also identity sentences but with the reversed relation of discriminability; reversing the elements does not make the sentences predicational.

We address this issue by reverting specificational sentences in Hebrew, which distinguishes predicational and specificational sentences in the choice of the copula. (51) is a specificational pseudocleft. Note that the copula is (neutral) pronZ.

(51) [ma še-dan menase li-mco ba-iton] ze [le-faxot štey ta'uyot]
what that-Dan tries to-find in-the-paper Z(n) at-least two mistakes
'What Dan is trying to find in the newspaper is at-least two mistakes.'

If we try to invert the element around the copula, we get an ungrammatical sentence with pronH, as in (52a), but a grammatical version with pronZ, as in (52b).

(52) a. *[le-faxot štey ta'uyot] hen [ma še-dan menase li-mco ba-iton]
at-least two mistakes H(f,pl) what that-Dan tries to-find in-the-paper
b. [le-faxot štey ta'uyot] ze [ma še-dan menase li-mco ba-iton]
at-least two mistakes Z(n) what that-Dan tries to-find in-the-paper
both: 'At least two mistakes are what Dan is trying to find in the newspaper.'

Note that the ungrammaticality of (52a) cannot be attributed to each of the phrases by themselves: as demonstrated in (53), both phrases can figure in the positions in a pronH sentence (these examples are from Heller 1999).

(53) a. **[le-faxot štey ta'uyot]** hen xamurot
at-least two mistakes H(f,pl) serious
'At least two mistakes are serious.'

b. ha-ta'ut še-šay asa hi **[ma še-dan menase li-mco ba-iton]**
the-mistake that-Shai made H what that-Dan tries to-find in-the-paper
'Shai's mistakes are what Dan is trying to find in the newspaper.'

In sum, we see that when a specificational sentence is inverted it yields an inverted specificational sentence and not a predicational one – this is expected if specificational sentences are identity sentences, but not if they are cases of inverse predication.

2.4. *Inversion or identity: summary*

This section has compared the discriminability analysis presented here with the inverse analysis of specificational sentences. Let us summarize the differences between the two.

First, the inverse analysis predicts that a predicational sentence can be inverted into a specificational one, modulo discourse constraints. The identity + discriminability analysis predicts this to be possible only if the predicational sentence expressed identity as a special case. We saw in section 2.1 that even if the predicate fits the discourse constraints, not all predicational sentences can be inverted. Note, importantly, that the discourse restriction proposed by Mikkelsen may be relevant for the identity analysis as well. As it stands, the identity + discriminability analysis cannot capture the difference between bare indefinites and indefinites containing discourse-old material in the pre-copular position of a specificational sentence.

Second, reversed specificational sentences are predicted to be predicational by the inverse analysis. We saw in the previous section that when inverting a specificational sentence in Hebrew we don't always get a predicational sentence. Rather, the inverted sentence has the same copula as a specificational sentence, i.e. pronZ. But the main problem with this prediction is that if these are indeed predicational sentences, they are not predicted to exhibit connectivity effects. (54-55) shows that reversed specificational

sentences do exhibit connectivity[4].

(54) *Principle A connectivity*
a. What John$_i$ is is proud of himself$_i$.
b. proud of himself$_i$ is what John$_i$ is.

(55) *Opacity connectivity*
a. What John is looking for is a pink giraffe.
b. A pink giraffe is what John is looking for.

Under the identity analysis, reversed specificational sentences are expected to exhibit connectivity effects, as these are a by-product of semantic equation[5]. In sum, while I have not presented an account of reversed specificational sentences, we see that the identity + discriminability analysis can potentially handle the issues raised by this type of sentences. The inverse analysis, on the other hand, does not distinguish these from predicational sentences and therefore does not have the tools to deal with this sentence type.

3. Towards a Typology of Copular Sentences

So far, this chapter has focused on specificational sentences and discussed predicational sentences in comparison. While these are the types most discussed in the literature, Higgins (1973) also distinguishes "pure" identity sentences and identificational sentences. I conclude the discussion in this dissertation by taking a look at his typology (56) exemplifies all four types.

---

[4] We saw in chapter 2 that den Dikken et al. (2000) observe that NPI connectivity is not found in reversed specificational pseudoclefts. Heycock & Kroch (2002) discuss further asymmetries in the connectivity pattern of the two: they show that reversed specificational sentences do not exhibit all the connectivity effects found in standard specificational sentences. Nonetheless, the fact that reversed specificational sentences exhibit connectivity beyond Bound Variable connectivity which is found in predicational sentences shows that they are not predicational.

[5] This does not necessarily concern all connectivity effects. It is only relevant for those that are a by-product of semantic equation. If some connectivity effects is affected by discourse factors, e.g. Principle B and Principle C connectivity, changing the linear order of the elements may create an environment that not longer licenses these effects.

(56) a. John is a doctor. *Predicational*
b. My next-door neighbor is Dan. *Specificational*
c. The evening star is the morning star. *Identity*
d. That is a kangaroo. *Identificational*

Higgins (1973) takes each type of copular sentence to have two slots, each hosting a certain kind of phrase – this is summarized in table 5.3.

Sentence Type	Pre-copular	Post-copular
Predicational	Referential	Predicational
Specificational	Superscriptional	Specificational
Identificational	Referential	Identificational
Identity	Referential	Referential

Table 5.3: Higgins' four-way copular typology

This typology is accompanied by a list of the expressions that can serve as these types (p. 264).

Type of expression	Referential	Superscriptional	Predicational	Specificational	Identificational
Deictic	+	-	-	+	+
Proper name	+	-	-	+	+
Definite NP	+	+	+	+	+
Indefinite NP	?-	?-	+	+	+
AP			+	+	-
VP, S			-	+	-

Table 5.4: Possible roles of different expressions

First, it should be pointed out that even when considering just the data presented by Higgins, these tables are not enough to predict which sentence will have what reading. To give just one example, consider the pre-copular phrases in identificational sentences. While Higgins says that this position should host either a demonstrative or an NP with a demonstrative determiner, the label for this position is "referential" which also includes proper name and definites. But the main problem with this characterization is that the two positions in a copular sentence are not independent – throughout this chapter we saw a

number of cases where a certain expression can be combined with one expression but not the other in order to yield a certain type of sentence – see again examples (1), (10) and (52a) above. In this section we consider whether and how Higgins' classification can be reduced in the light of the discriminability analysis. But first we consider Mikkelsen's (2004a,b) attempt to reduce the typology. To the best of my knowledge, this is the only such attempt in the literature.

3.1. *Identificational sentences*

Higgins (1973, p. 238) presents the following examples as identificational sentence. While the pre-copular phrase must be a demonstrative pronoun or a headed description with a demonstrative determiner (hereafter – demonstrative NP), the post-copular phrase can vary: it can be a proper name, as in (57), a definite, as in (58), or an indefinite, as in (59-60).

(57) a. That is Boston.
b. That place is Boston.

(58) a. That is the house I mentioned.
b. This house is the house I mentioned.

(59) a. That is aluminum.
b. That stuff is aluminum.

(60) a. That is a kangaroo.
b. That animal is a kangaroo.

According to Higgins, identificational sentences are "typically used for teaching the names of people and things" (Higgins, p. 237).

Mikkelsen (2004a,b) has proposed reducing Higgins' typology to a three-way typology by splitting the identificational class and integrating the two subtypes into other already existing copular types. In particular, Mikkelsen proposes distinguishing

identificational sentences in which the pre-copular phrase is a demonstrative from identificational sentences in which the pre-copular phrase is an NP with a demonstrative determiner.

(61) a. That is Susan. (Mikkelsen 2004b)
b. That woman is Susan.

In support of her claim, Mikkelsen presents two tests on which the pre-copular phrases in (61) behave differently[6]. (62) presents the test of pronominalization: in a tag question[7], the pre-copular phrase in (61a) can be referred to using *it* while the pre-copular phrase in (61b) can only be referred to using the personal pronoun *she*.

(62) *Pronominalization*
a. That is Susan, isn't (it/*she)?
b. That woman is Susan, isn't (*it/she)?

Mikkelsen assumes that the pronoun *she* refers to individuals of type e while the pronoun *it* refers to properties at type <e,t>. Accordingly, the contrast in (62) indicates that the pre-copular expression in (61a) denotes a property while the pre-copular expression in (61b) denotes an individual. The same conclusion is reached based on the modification test in (63). In (63a) the bare demonstrative cannot be modified with a non-restrictive relative clause introduced by *who*, while the demonstrative in (63b) can.

(63) *Non-restrictive Modification*
a. *That, who everybody can see clearly, is Susan.
b. That woman, who everybody can see clearly, is Susan.

---

[6] Mikkelsen present a third argument which concerns the form of the determiner is Danish in both kinds of sentences. Unfortunately, I had no access to a Danish consultant who could help me with extending the discussion in this chapter to Danish.

[7] The tag question with *it* is of course grammatical under the irrelevant reading where *it* refers back to the sentences as a whole.

Since non-restrictive modification is expected to be available only for modifying individuals, Mikklsen concludes that the pre-copular phrase in (61a) does not denote an individual.

Based on these data, Mikkelsen concludes that (61a), and other identificational sentences with a bare demonstrative in the pre-copular phrase, are composed of the types [<e,t> *is* e] and hence are (elided) specificational sentences – recall that Mikkelsen is a proponent of the inverse analysis for specificational sentences. (61b) and other identificational sentences with a demonstrative NP in the pre-copular phrase, are determined to be composed of the types [e *is* e] which makes them identity sentences.

However, if we examine the behavior of these sentences on Higgins' tests, we find that all the identificational sentences are actually similar to predicational sentences. This contrast with Mikkelsen's proposal which predicts that the sentences with the bare demonstrative will pattern like specificational sentences. For the sentences with the demonstrative NP no prediction is possible as we do not know how "pure" identity sentences pattern on these tests. It should be pointed out that Higgins himself does not discuss identificational sentences in the context of these tests.

We conduct the tests on four different identificational sentences: two with a bare demonstrative in (64) and (66) and two with a demonstrative in (65) and (67). Note that the first two are [+human] and the other two are [-human]. The (a) examples are the sentences themselves. The (b) examples test coordination of predicates, the (c) examples test deleting the post-copular phrase in a coordination and the (d) examples test deleting the copula in a coordination.

(64) a. That is Susan.
b. *That is Susan and is very cute.
c. That is Susan and that is too.
d. That is Susan and that … Michelle.

(65) a. That woman is Susan.
b. That woman is Susan and is very cute.
c. That woman is Susan and that woman is too.
d. That woman is Susan and that woman … Michelle.

(66) a. That is a kangaroo.
b. That is a kangaroo and is very cute.
c. That is a kangaroo and that is too.
d. That is a kangaroo and that … a koala.

(67) a. That animal is a kangaroo.
b. That animal is a kangaroo and is very cute.
c. That animal is a kangaroo and that animal is too.
d. That animal is a kangaroo and that animal … a koala.

First, note that all four sentences allow deleting the copula in a coordination – this is a characteristic of predicational sentences. Next, note that all four sentences allow deleting the post-copular phrase in a coordination – again, a characteristics of predicational sentences. It should be pointed out that for (64) and (65) this requires a context of pointing at pictures rather at the woman directly – this is the only way to make sense of the coordination. Finally, note that coordinating an adjectival predicate is not possible in only in (64b). I thus conclude that identificational sentences are in fact instances of predicational sentences. That is, both the bare demonstrative and the demonstrative NP pick out an entity.

I would like to propose that this ungrammaticality of (64b) is because the entity picked out by the bare demonstrative bears no [±human] feature in the model, and as a result a [+human] predicate is inappropriate. This may seem strange because the entity corresponds to a [+human] entity in the world. Nonetheless, I propose that picking it out

by ostension only renders the entity in the model specified only for special properties. Note that this can also account for the use of an *it* pronoun in the tag question in (62a) and the ban on a *who* relative clause in (63a). Once the information in the identificational sentence has been added into the context, the entity is [+human] and it cannot be referred to using *it*.

Going back to identificational sentences, if they pattern with predicational ones on Higgins' tests, why did Higgins take them to be a separate class? In an approach which classifies sentence into types based on the expressions that can occur in each of the two positions, predicational and identificational sentences look very different. For one, the former allows for AP predicates which are banned from the latter. Additionally, Higgins assumed that identificational but not predicational sentences allow for proper names in the post-copular position (but see again section 1.3). My approach instead takes the relation between the two expressions to determine the type of the sentence, thereby allowing us to see the similarities between these sentence types.

In particular, I propose that identificational sentences are a special case of predicational sentences. Intuitively, these are predicational sentences in which the post-copular phrase is an essential property that is assumed to be new information, so the entity denoted by the pre-copular phrase has to be picked out by an expression that does not presuppose this property. Note that in the light of this analysis we may wish to classify the ambiguous cat examples in (13-14) as ambiguous between identificational and specificational and not between predicational and specificational as proposed by Higgins.

3.2. *Pure identity?*

I have argued in the previous section that identificational sentences are in fact a special case of predicational sentences – this allows us to reduce Higgins' four-way typology into a three-way typology distinguishing predicational, specificational and identity sentences. But we have not yet discussed identity sentences. For Higgins, identity sentences equate two entities – he adopts this type of sentences from the philosophical tradition without much discussion. Some examples are given in (68).

(68) a. The evening star is the morning star.
b. Cicero is Tully.

I have proposed in this chapter that specificational sentences are identity sentences in which the post-copular phrase is more discriminate than the pre-copular phrase. That is, the sentence identifies an entity that is less familiar with an entity that is more familiar. I have further proposed that the opposite discriminability relation is found in the more marked constriction of reversed specificational sentences. In addition, we have seen sentences that express identity as a special case of predication. What Higgins calls identity sentences raises the question of whether pure identity exists.

If pure identity sentences exist, these should be sentences in which the two expressions are of equal discriminability, i.e. sentences where one expression does not reflect more familiarity than the other. Due to our limited understanding of discriminability, I do not know at this point how to exam this claim. However, we have seen in this chapter that discriminability should be relativized to the knowledge state of the interlocutors in a specific discourse (see again section 1.3), so it seems unlikely that there will be two expressions that render the entity at the exact same level of familiarity. If the context always renders one expression more discriminate than the other, identity

sentences will have either rising or decreasing discriminability and "pure" identity will not exist. I leave this issue for future research.

4. Conclusions

This chapter used the notion of discriminability introduced in the previous chapter to study how specificational sentences are different from other sentences that express identity. We saw that specificational sentences are special in that the post-copular phrase must be more discriminate than the pre-copular phrase. We can thus view specificational sentences as identifying a less familiar entity with a more familiar entity. For predicational sentences, we saw that discriminability is only relevant for those predicational sentences that express identity as a special case. Finally, we take reversed specificational sentences to be identity sentences with the reversed discriminability of specificational ones. I have proposed that this is the marked discriminability relation in identity sentences.

In different points in this chapter, we saw that nominal predicates can be of different kinds. First, there are nominal predicates which do not encode the sortal concept of the noun but rather a property of it, similar to adjectival predicates. But sortal concepts are also not all equal: depending on the nature of the entity, some sortal concept may be more inherent than another. These distinctions require appropriate representation. Standardly, all properties are represented as predicates. The data presented in this chapter suggests that different representations are required for the different kinds of nominal predicates – see Gupta 1980 for a proposal distinguishing sortal from non-sortal predicates. Note that this issue goes beyond the analysis of copular sentences.

Going back to copular sentences, we can now answer the questions raised in the first chapter.

- Can specificational sentences be defined as copular sentences that exhibit connectivity? No. We saw in chapter 1 that the connectivity effect of Bound Variable is also found in predicational sentences. Therefore, if a sentence exhibits connectivity, it is not entailed that it is specificational. In addition, we saw in chapter 3 that agreeing pronZ pseudoclefts in Hebrew exhibit only some connectivity effects. Therefore, if a sentence is specificational, it is not entailed that it would exhibit connectivity.
- Can specification be analyzed in such a way that will also account for connectivity? Analyzing specificational sentences as identity sentences account for connectivity, but they are other identity sentences which are not specificational.
- Are these two phenomena independent? Yes. Connectivity is a by-product of certain syntactic and semantic combinatorics, whereas specification concerns some pragmatic relation between the phrases around the copula. The two are found in specificational sentences because of the identity relation.

The discriminability analysis also sheds new light on the copular sentences and allows reducing Higgins' four-way typology to a three-way typology. In particular, I have argued that identificational sentences are a special case of predicational sentences, distinguishing predicational, specificational and pure identity sentences. It is my hope that a better understanding of discriminability will allow reducing the typology further. If there are no "pure" identity sentences, and all identity sentences will have either rising or decreasing discriminability, the distinction between identity sentences and specificational

sentences will not be necessary. Instead, we will have two possible semantic relations in copular sentences – predication and identity – and further nuances would be a result of discriminability – rising or decreasing.

REFERENCES

Akmajian, Andreas: 1970, 'On Deriving Cleft Sentences from Pseudocleft Sentences', *Linguistic Inquiry* **1**, 140-168.

Alexiadou, Artemis and Anastasia Giannakidou: 1998, 'Specificational Pseudoclefts and the Semantics of Lists', *ZAS Working Papers on Linguistics 10*, pp. 1-21.

Ariel, Mira: 1988, 'Referring and Accessibility', *Journal of Linguistics* **24**, 67-87.

Bach, Emmon and Barbara H. Partee: 1980, 'Anaphora and Semantic Structure', in J. Kreiman and A. Ojeda (eds.), *CLS Parasession on Anaphora*, pp. 7–28, Chicago Linguistic Society, Chicago.

Baker, Carl L.: 1968, *Indirect questions in English*, Ph.D. dissertation, University of Illinois.

Berman, Ruth and Alexander Grosu: 1976, 'Aspects of the Copula in Modern Hebrew', in P. Cole (ed.), *Studies in Modern Hebrew Syntax and Semantics*, pp. 265-285, North Holland, Amsterdam.

Birner, Betty. J.: 1996, *The Discourse Function of Inversion in English*, New York: Garland.

Bresnan, Joan J. and Jane Grimshaw: 1978, 'The Syntax of Free Relatives in English', *Linguistic Inquiry* **9**, 331–391.

Brown, Roger: 1965, 'How shall things be called?', in R. C. Oldfield and J. C. Marshall (eds.), *Language*, pp. 82-92, Harmondsworth, Penguin.

Caponigro, Ivano: 2003, *Free Not to Ask: On the Semantics of Free Relatives and Wh-Words Cross-linguistically*, Ph.D. dissertation. University of California, Los Angeles.

Caponigro, Ivano and Daphna Heller: to appear, 'The Non-Concealed Nature of Free Relatives: Implications for Connectivity', in C. Barker and P. Jacobson (eds.), *Direct Compositionality*, Oxford University Press.

Cecchetto, Carlo: 2000, 'Connectivity and Anti-Connectivity in Pseudoclefts', in M. Hirotani (ed.), *Proceedings of NELS 30*, pp. 137-151, GLSA, University of Massachusetts, Amherst.

——————— 2001, 'Syntactic or Semantic Reconstruction? Evidence from Pseudoclefts and Clitic Left Dislocation', in C. Cecchetto, G. Chierchia and M. T. Guasti (eds.), pp. 90–144, *Semantic Interfaces*, CSLI, Stanford, California.

Chierchia, Gennaro: 1984, *Topics in the Syntax and Semantics of Infinitives and Gerunds*, Ph.D. dissertation, University of Massachusetts, Amherst.

——————— 1993, 'Questions with Quantifiers', *Natural Language Semantics* **1**, 181-234.

Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.

——————— 1986, *Knowledge of Language: Its Nature, Origin, and Use*. New York: Praeger.

Clark, Herbert H. and Catherine Marshall: 1981, 'Definite Reference and Mutual Knowledge', in A. K. Joshi, B. L. Webber and I. A. Sag (eds.), *Elements of discourse understanding*, pp. 10-63, NY: Cambridge University Press.

Cruse, D. Alan: 1977, 'The Pragmatics of Lexical Specificity ', *Journal of Linguistics* **13**, 153-164.

Dahl, Östen: 1981, 'Binding Relations in Dislocated Constituents', talk presented at the Sloan Workshop on Alternatives to Transformational Grammar, Stanford University.

Danon, Gabi: 2001, 'The Hebrew Object Marker as a Type-shifting Operator', in E. Aim, K. Gerdes and H. Yoo (eds.), *Proceedings of the 6th Doctorate Meeting in Linguistics*, pp. 41-46, Paris: Université Paris 7.

Dayal, Veneeta: 1997, 'Free Relatives and *–ever*: Identity and Free Choice Readings', in A. Lawson and E. Cho (eds.), *Proceedings of SALT VII*, pp. 99-116, CLC Publications, Ithaca NY.

den Dikken, Marcel, André Meinunger, and Chris Wilder: 2000, 'Pseudoclefts and ellipsis', *Studia Linguistica* **54**, 41-89.

Donnellan, Keith S.: 1966, 'Reference and Definite Descriptions', first appeared in *The Philosophical Review* **77**, 281-304.

Dor, Daniel: 1992, 'Towards a Semantic Account of Concealed Questions', in M. Bernstein (ed.), *Proceedings of the Ninth Eastern States Conference on Linguistics (ESCOL '92)*, pp. 56-67, Ithaca, NY: Cornell University.

Doron, Edit: 1983, *Verbless Predicates in Hebrew*, Ph.D. dissertation, University of Texas at Austin.

Dowty, David and Pauline Jacobson: 1989, 'Agreement as a Semantic Phenomenon.', in J. Powers and K. de Jong (eds.) *Proceedings of the Fifth Eastern States Conference on Linguistics (ESCOL '88)*, pp. 95-108, Columbus: Ohio State University.

Engdahl, Elisabeth: 1986, *Constituent Question*s, Kluwer, Dordrecht.

von Fintel, Kai: 2000, 'Whatever', in B. Jackson and T. Matthews (eds.), *Proceedings of SALT X*, pp. 27–40, Ithaca, NY: CLC Publications.

Greenberg, Yael: 1994, *Hebrew Nominal Sentences and the Nature of the Stage-Individual Level Distinction*, M.A. Thesis, Bar-Ilan University.

——————— 1998, 'An Overt Syntactic Marker for Genericity in Hebrew', in S. Rothstein (ed.), *Events and Grammar*, pp. 125-143, Kluwer Academic Press.

Grice, Herbert Paul: 1975, 'Logic and conversation', in P. Cole and J.L. Morgan (eds.), *Speech Acts*, pp. 41–58, New York: Academic Press.

Grimshaw, Jane: 1979, 'Complement selection and the lexicon', *Linguistic Inquiry* **10**, 279-326.

Grodzinsky, Yosef and Tanya Reinhart: 1993, 'The Innateness of Binding and Coreference', *Linguistic Inquiry* **2**4, 69–101.

Groenendijk, Jeroen and Martin Stokhof: 1984, *Studies on the Semantics of Questions and the Pragmatics of Answer*s, Academisch Proefschrift, Amsterdam.

——————— 1990, 'Partitioning Logical Space', annotated handout, Second European Summerschool in Logic, Language and Information, Leuven, August 1990.

——————— 1997, 'Questions', in J. van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, pp. 1055–1124, Cambridge, Massachusetts, MIT Press.

Groenendijk, Jeroen, Martin Stokhof and Frank Veltman: 1996, 'Coreference and Modality', in S. Lappin (ed.) *Handbook of Contemporary Semantic Theory*, pp. 179-213, Oxford, Blackwell.

Groos, Anneke and Henk van Riemsdijk: 1981, 'Matching effects with Free Relatives: A Parameter of Core Grammar', in A. Belleti, L. Brandi and L. Rizzi (eds.) *Theories of Markedness in Generative Grammar*, pp. 171-216, Scuola Normale Superiore di Pisa, Pisa.

Grosu, Alexander: 1994, *Three Studies in Locality and Case*, Routledge, London / New York.

Gundel, Jeanette, Nancy Hedberg and Ron Zacharski: 1993, 'Cognitive status and the form of referring expressions in discourse', *Language* **69**, 274-307.

Gupta, Anil: 1980, *The Logic of Common Nouns: An Investigation in Quantified Modal Logic*, Yale University Press, New Haven.

Halvorsen, Per-Kristian: 1978, *The Syntax and Semantics of Cleft Constructions*, Ph.D. dissertation, University of Texas, Austin.

Hamblin, Charles Leonard: 1973, 'Questions in Montague English', *Foundations of Language* **10**, 41-53.

Heggie, Laurie : 1988, *The Syntax of Copular Sentences*, Ph.D. dissertation, University of Southern California.

Heim, Irene: 1979, 'Concealed Questions', in R. Bäuerle, U. Egli and A. von Stechow (eds.), *Semantics from Different Points of View*, pp. 51-60, Berlin: Springer-Verlag.

——————— 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts, Amherst.

Heller, Daphna: 1999, *The Syntax and Semantics of Specificational Pseudoclefts in Hebrew*, MA thesis, Tel Aviv University.

——————— 2002, 'On the Relation of Connectivity and Specificational Pseudoclefts', *Natural Language Semantics* **10**, 243-284.

Heycock, Caroline and Anthony Kroch: 1997, 'Inversion and Equation in Copular Sentences', paper presented at the workshop on (Pseudo)clefts, ZAS, Berlin.

——————— 1999, 'Pseudoclefts Connectivity: Implications for the LF Interface Level', *Linguistic Inquiry* **30**, 365-397.

——————— 2002, 'Topic, Focus, and Syntactic Representations', in L. Mikkelsen and C. Potts (eds.) *Proceedings West Coast Conference on Formal Linguistics 21*, pp. 141-165, Santa Cruz.

Higgins, F. Roger: 1973, *The Pseudocleft Construction in English*, Ph.D. dissertation, MIT, Cambridge, MA [version of 1979 published by Garland, New York].

Hornstein, Norbert: 1984, *Logic as Grammar*, MIT Press, Cambridge.

Iatridou, Sabine and Spyridoula Varlokosta: 1998, 'Pseudoclefts Crosslinguistically', *Natural Language Semantics* **6**, 3-28.

Izvorski, Roumyana: 1997, 'On the Type of *be* and the Nature of the *wh*-clause in Specificational Pseudoclefts', paper given at the Workshop on the Syntax and Semantics of (Pseudo-)clefts, ZAS, Berlin, November 1997.

——————— 2000, 'Free Adjunct Free Relatives', in R. Billerey and B. D. Lillehaugen (eds.), *Proceedings of the 19th West Coast Conference on Formal Linguistics*, pp. 232-245.

Jacobson, Pauline: 1994, 'Binding Connectivity in Copular Sentences', in M. Harvey and L. Santelmann (eds.), *Proceedings of SALT IV*, pp. 161–178, Cornell University, Ithaca.

——————— 1995, 'On the Quantificational Force of English Free Relatives', in E. Bach, E. Jelinek, A. Kratzer and B. Partee (eds.), *Quantification in Natural Language*, pp. 451-486, Kluwer Academic Publishers, Dordrecht.

——————— 1999, 'Towards a Variable-Free Semantics', *Linguistics and Philosophy* **22**, 117-184.

Kadmon, Nirit: 2001, *Formal Pragmatics: Semantics, Pragmatics, Presupposition, and Focus*. Oxford: Blackwell Publishers.

Kamp, Hans: 1981, 'A Theory of Truth and Semantic Representation', in J. Groenendijk, T. Janssen and M. Stokhof (eds.), *Formal Methods in the Study of Language: Proceedings of the Third Amsterdam Colloquium*, pp. 277-322, Mathematical Centre Tracts, Amsterdam.

Karttunen, Laurie: 1977, 'Syntax and Semantic of Questions', *Linguistics and Philosophy* **1**, 3-44.

Mikkelsen, Line: 2004a, 'Reexamining Higgins' Taxonomy: A Split in the Identificational Class', LSA Annual Meeting. Boston, January 10.

——————— 2004b, *Specifying Who: On the Structure, Meaning and Use of Specificational Copular Clauses*, Ph.D. dissertation, University of California, Santa Cruz.

Montague Richard: 1973, 'The Proper Treatment of Quantification in Ordinary English', in K. J. J. Hintikka, J. M. E. Moravcsik and P. Suppes (eds.), *Approaches to Natural Language,* pp. 221-242, Reidel, Dordrecht.

Moro, Andrea: 1997, *The Raising of Predicates: Predicative Noun Phrases and the Theory of Clause Structure*, Cambridge: Cambridge University Press.

Nathan, Lance: forthcoming, *The Interpretation of Concealed Questions*, Ph.D. dissertation, MIT.

Partee, Barbara H.: 1986, 'Ambiguous Pseudoclefts with Unambiguous BE', in S. Berman, J.-W. Choe, and J. McDonough (eds.), *Proceedings of NELS 16*, pp. 354–366, GLSA, University of Massachusetts, Amherst.

——————— 1987, 'Noun Phrase Interpretation and Type Shifting Principles', in J. Groenendijk, D. de Jongh and M. Stokhof (eds.) *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Foris, Dortrecht, pp. 115-143.

Peters, Stanley and Emmon Bach: 1968, 'Pseudo-cleft sentences', unpublished ms., University of Texas, Austin.

Pollard, Carl, and Ivan A. Sag: 1983, 'Reflexives and Reciprocals in English: An Alternative to the Binding Theory', in M. Barlow, D. Flickinger, and M. Westcoat (eds.), *Proceedings of the 2nd West Coast Conference on Formal Linguistics*, pp. Stanford Linguistics Association, 189-203, Stanford, CA.

——————— 1992, 'Anaphors in English and the Scope of the Binding Theory', *Linguistic Inquiry* **23**, 261–303.

Prince, Ellen F.: 1992, 'The ZPG Letter: Subjects, Definiteness, and Information Status', in W. Mann and S. Thompson (eds.), *Discourse Description: Diverse Linguistic Analyses of A Fund-raising Text*, pp. 295-325, Amsterdam: John Benjamins.

Quer, Josep: 1998, *Mood at the Interface*. The Hague: Holland Academic Graphics.

Rapoport, Tova Rebecca: 1987, *Copular, Nominal and Small Clauses: a Study of Israeli Hebrew*, Ph.D. dissertation, MIT, Cambridge, MA.

Reinhart, Tanya: 1997, 'Quantifier Scope - How Labor is Divided between QR and Choice Functions', *Linguistics and Philosophy* **20**, 335-397.

Reinhart, Tanya and Eric Reuland: 1993, 'Reflexivity', *Linguistic Inquiry* **24**, 657–720.

Romero, Maribel: 2003, 'Concealed Questions and Specificational Subjects', ms. University of Pennsylvania.

——————— in press, 'Concealed Questions and Specificational Subjects', *Linguistics and Philosophy*.

Ross, J.R.: 1972, 'Act', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, pp.70-126, Reidel, Dortrecht.

——————— 1997, 'That is the Question', paper presented at the University of Pennsylvania.

Rothstein, Susan: 1995, 'Small Clauses and Copular Constructions', in A. Cardinaletti and M. T. Guasti (eds.), *Small Clauses*, pp. 27-48, New York: Academic Press.

——————— 2000, *Predicates and Their Subjects*, Kluwer Academic Press.

Schlenker, Philippe: 2003, 'Clausal equations (A note on the connectivity problem)', *Natural Language and Linguistic Theory* **21**, 157–214.

Sharvit, Yael: 1997, *The syntax and Semantics of Functional Relative Clauses*, Ph.D. dissertation, Rutgers University.

——————— 1999, 'Connectivity in Specificational Sentences', *Natural Language Semantics* **7**, pp. 299-339.

Sichel, Ivy: 1997, 'Two Pronominal Copulas and the Syntax of Hebrew Nonverbal Sentences', *Proceedings of the TLS Conference on Predication*.

Stalnaker, Robert: 1978. 'Assertion', in P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*, pp. 315-332, Academic Press.

Szabolcsi, Anna: 1987, 'Comparative Superlatives', in N. Fukui, T. R. Rapoport and E. Sagey (eds.), *Papers in Theoretical Linguistics*, pp. 245–265. MIT Working Papers in Linguistics, Cambridge, Mass.

Tredinnick, Victoria: 2005, *On the Semantics of Free Relatives with '-ever'*, Ph.D. dissertation, University of Pennsylvania

Vikner, Carl and Per Anker Jensen: 2002, 'A Semantic Analysis of the English Genitive. Interaction of Lexical and Formal Semantics', *Studia Linguistica* **56(2)**, 191-226.

Williams, Edwin: 1983, 'Semantic vs. Syntactic Categories', *Linguistics and Philosophy* **6**, 423- 446.

——————— 1994: *Thematic Structure in Syntax* (Linguistic Inquiry Monograph 23), MIT Press, Cambridge, Mass.

Zimmermann, Thomas Ede: 1993, 'On the Proper Treatment of Opacity in Certain Verbs', *Natural Language Semantics* **1**, 149–179.