



Durham E-Theses

Predication and identity in copular sentences

WOODARD, ANDREW

How to cite:

WOODARD, ANDREW (2018) *Predication and identity in copular sentences*, Durham theses, Durham University. Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/12890/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Predication and identity in copular sentences

Department of philosophy, Durham University
Submitted for the degree of doctor of philosophy
by

Andrew Woodard

September 10, 2018

Abstract

There are different types of copular sentence. *Cicero is tall* does not mean the same thing as *Cicero is Tully*. The former is typically called a predication and the latter an identity. But where does this difference come from (and the difference between these sentences and *Cicero is a Roman statesman*, *This is Cicero*, and *The culprit was Cicero* for that matter)? This thesis is an attempt to bring the combined forces of modern linguistics and philosophy together to understand how we make meaningfully distinct sentences. It is useful to focus on copular sentences for this task because they are the minimally-sized sentences in many of the world's languages. A debate in linguistics has persisted for some years about the status of sentences like *The culprit was Cicero*, in particular, in terms of whether they should be aligned with predications or identities. The linguistic evidence points in different directions. I think there are conceptual clarifications that could elucidate the terms of this debate.

I start by investigating the obvious logical starting point: logic. Can copular sentences really be exhaustively specified as logical predications or identities in the first place? What about syntax? Does it have a problem-free definition of predication that might serve as a way of distinguishing meanings? I answer in the negative in both cases. From there the copula itself is studied, and I argue that it isn't any kind of real lexical verb that can be semantically ambiguous (like the verb *bank* can in *I'm going down a driveway banked by boulders and wildflowers* vs. *I banked the cheque then went for lunch*). I then give some examples of the sort of thing linguists have shown copulas actually can do. I argue that distinguishing copular sentence meanings is not something that can be truly *explained* in logical or syntactic terms; the best we can hope for is to *describe* the different uses of the biological capacity for language, to talk of 'functions' of copular sentences and their components. I argue for an approach to a full description of copular sentences that is based mainly on use and information structure properties of the flankers of the copula. Different permutations of these give rise to different sentence types, and thus to a more pluralistic copular taxonomy. Syntax has a foundational role, but it does not serve to discriminate copular sentence meanings. Copular sentences are argued to be *uniformly* syntactic predications.

Overall, then, I argue that the taxonomy of copular sentences cannot be explained in terms of logical predication or logical identity. Rather, we can describe (not explain) the distinct meaning types of copular sentences in terms of the number of different licit permutations of flanker properties, where these are mainly use and information structure properties.

Dedication

In memory of my mother, Carolyn Seabrook-Smith Woodard, who would have loved to see me finally finish this!

Statement of copyright

The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Acknowledgements

I originally came to Durham to work with Wolfram Hinzen, whose views on language had fascinated me when I first encountered them some years before. Supervisions with Wolfram were always a rich intellectual experience. When Wolfram moved to Barcelona, Sara Uckelman stepped in and it has been a great help throughout the writing-up period to have her to bounce ideas off and make sure my writing is clear. I also thank my secondary supervisor, Ben Smith, who provided me with extremely useful comments on the final draft.

It was excellent to be a part of the Language and Mind research group (now sadly disbanded) and I thank the members of that group for interesting and useful presentations and discussions: Uli Reichard, James Miller, Alex Malt, David Kirkby, Tom Hughes, and Pallavi Worah. Uli, in particular, gave me very useful comments on parts of this thesis. Txuss Martin has been a great friend and source of enthusiasm and interesting discussion. Will Jones has also been a great friend and I thank him for the countless useful discussions about matters pertaining to language and the brain.

Elsewhere in the philosophy department I benefited from discussions with Jonathan Lowe before he died so sadly and prematurely. The Mind, Language, and Metaphysics meetings were always constructive as were the audiences at the Eidos talks I gave. I thank the department itself for departmental scholarships that provided me with financial assistance. Other friends in the department provided moral support, especially Tom Rossetter and Jamie Taylor.

I also wish to thank several scholars for discussions and/or providing me with books and articles that helped me write this thesis: Isabelle Roy, Jutta Hartmann, Amanda Patten, Claire Beyssade, Thom Scott-Phillips, Mariann Slíz, John Collins, and Matthew Tugby.

Marine Clément supported me for a long time, which I thank her for. During the course of this thesis, both my parents and a beloved aunt all died prematurely. The loss of my mother, in particular, was hard; she was certainly my most unflinching supporter and I miss her every day.

Contents

Abbreviations and notation	10
List of figures and tables	12
1 Introduction: the taxonomy of copular sentences	13
1.1 Aim of this thesis	13
1.2 The basic data	14
1.2.1 Debates about right categorisation	16
1.2.2 Extending Higgins’s four categories to six, Roy (2013)	21
1.3 An eight-ways ambiguous example	31
1.4 A guide to the rest of the thesis	33
2 Can the logical or syntactic notions of predication help explain the taxonomy of copular sentences?	38
2.1 Introduction	38
2.1.1 Some philosophical approaches to the meaning of predication	42
2.2 Predication in the work of Frege	45
2.2.1 The meaning of a predicate is a concept	45
2.2.2 Concepts as functions	46
2.2.3 How to determine the function for a given sentence	49
2.2.4 Functions as sets	52
2.2.5 Predicate modification and set intersection	53
2.2.6 Intensionality	54
2.2.7 Identity as grammatical predication in Frege	55
2.2.8 Conclusions	57
2.3 Functions in linguistics, mainly syntax	58
2.3.1 Theta roles	59
2.3.2 Internal and external arguments	61

2.3.3	Subjects as external theta roles	67
2.3.4	Might subjects simply be the same as external arguments?	70
2.3.5	Eliminating predication?	72
2.3.6	Type-shifting	75
2.3.7	Merge as a function	76
2.4	Conclusion	78
3	Can the philosophical notion of identity help describe or explain the taxonomy of copular sentences?	81
3.1	Introduction	81
3.2	‘Pure’ definitions of identity from logic and metaphysics	84
3.2.1	Putative types of identity	84
3.2.2	Leibniz’s law	89
3.2.3	Eliminating identity and identity as indefinable	94
3.3	Meanings of identity statements	96
3.3.1	Frege’s ‘2-signs’ theory of identity statements	99
3.3.2	Frege’s ‘On Sense and Reference’ view of identity statements	104
3.3.3	Giving abbreviations	107
3.3.4	Overcoming an expectation of non-identity	108
3.3.5	Aspects of identity statements	110
3.4	Conclusion	112
4	Is the copula meaningful?	119
4.1	Introduction	119
4.2	The factual discovery of nominal sentences	122
4.2.1	More details about copula variation cross-linguistically	124
4.2.2	Semantic differences between nominal sentences and copular sentences according to Benveniste	127
4.2.3	The presence/absence of the copula in predicational vs. equative copular sentences	130
4.3	Lexical vs. functional categories	131
4.3.1	Copulas are inflection-like in many languages	133
4.3.2	Copulas are, or are part of, a closed class	134
4.3.3	Copulas are semantically functional-like	135
4.3.4	Copulas are distributed like auxiliaries	136
4.3.5	Copulas phonetically reduce like functional words, except in equatives	137

4.4	Is the verbal copula even a real verb?	137
4.4.1	Does the copula assign theta roles?	138
4.4.2	The copula as a non-case-assigning raising verb	141
4.5	Conclusion	146
5	What the copula does: two case studies	147
5.1	Introduction	147
5.2	The copula in English: a case study	148
5.2.1	The evidence for different meanings when the matrix and embedded subjects are quantified terms	149
5.2.2	Syntactic analysis of the different scope relations—‘small clause re- structuring’	151
5.3	<i>Ser/estar</i> in Spanish: a case study	163
5.4	Conclusion	168
6	Functions	170
6.1	What am I trying to explain?	170
6.2	Functions in Higgins	171
6.3	Prerequisites to giving the function of a sentence or parts of a sentence . . .	175
6.3.1	What gets used?	175
6.3.2	I-language can’t simply be used <i>tout court</i>	178
6.4	The way in which I-language <i>can</i> be used and the function of a sentence . .	183
6.4.1	A function for ostensive-inferential communication	184
6.4.2	Reconciling I-language with the function of ostensive-inferential com- munication	185
6.5	We can only describe, not explain, the taxonomy, and some ‘conceptual diffi- culties’	187
6.5.1	The relation between sentence and flanker function: two possible ways to characterise?	189
6.6	Conclusion	195
7	Pieces of the puzzle	197
7.1	Introduction	197
7.2	What logic gives us	198
7.3	A note about metaphysics and property ascription	199
7.3.1	Returning to language	202
7.3.2	Identificationals	203

7.4	What (else) philosophy gives us	204
7.5	What syntax gives us—syntactic predication	205
7.5.1	Merge	206
7.5.2	Phases	208
7.5.3	Merge and phasal coincidence—a redundancy	209
7.5.4	Predication Phrase?	210
7.5.5	Is there anything else to say about syntax?	212
7.6	Information Structure	215
7.6.1	A need for something else	215
7.6.2	Back to Frege	216
7.6.3	Basics of information structure	216
7.7	Multiple taxonomies and pieces of the puzzle	220
7.7.1	Multiple taxonomies and a ‘full description’	220
7.7.2	Pieces of the puzzle	221
7.8	Conclusion	224
8	Putting it all together: a case study	225
8.1	Introduction	225
8.2	The source of two readings—predicational and specificational	228
8.2.1	Focus	228
8.2.2	Referential vs. attributive	230
8.3	A counter-example to the view that (1) and (2) are semantically equivalent	231
8.3.1	A ‘well-noted maximality difference’ with pre- vs. post-copula possessives	232
8.3.2	Hartmann (forthcoming)	233
8.4	An alternative analysis	241
8.4.1	Exhaustivity	241
8.4.2	The overall picture	244
8.4.3	Alleviating stipulations	246
8.4.4	Extension to pseudoclefts	247
8.4.5	A more pluralistic taxonomy	249
8.4.6	Returning to Higgins’s eight-ways ambiguous example	253
8.4.7	Conclusions in the light of the comparison with Higgins	255
8.5	Arregi et al. (2018)—Arguments that (1) and (2) are not equivalent in their formal semantics	257
8.6	Conclusion	260

9 Conclusion	262
9.1 Summary	262
9.2 Desiderata	264
9.2.1 A relation between the flankers? And what is predication?	264
9.2.2 What is identity?	265
9.2.3 Re-analysing evidence for identity and predication	266
References	268

Abbreviations and notation

λ lambda operator

\exists existential quantifier

\forall universal quantifier

* ungrammatical sequence

infelicitous sequence

(...) optional sequence

(*...) sequence must be omitted

*(...) sequence cannot be omitted

3sg third person singular agreement

AGR agreement marker

AP adjective phrase

AuxP auxiliary phrase

CP complementiser phrase

COP copula

CIP classifier phrase

Det determiner

DP determiner phrase (see footnote 13, chapter 2)

FEM feminine

i-level individual-level predicate, see 1.2.2

ι iota operator

IP inflection phrase

NP noun phrase

NumP number phrase

PL plural

PP prepositional phrase

PredP predicate phrase

PRES present tense

RP relator phrase—see section 5.2.2.3.2

S sentence

SC small clause

s-level stage-level predicate, see 1.2.2

Spec,XP the specifier of XP

t trace

TP tense phrase

UG Universal Grammar

vP ‘little *v* phrase’—a ‘light verb’ functional head (see section 2.3.4; ‘*v**P’ is sometimes used to specifically refer to the functional head that introduces transitive verb phrases, see Chomsky (2008))

VN verbal noun

VP verb phrase

X' an intermediate phrasal projection (see footnote 9, chapter 2)

XP a phrase of any lexical or functional category

List of Figures

3.1	Illustration of Frege’s example, Frege (1892b)	102
6.1	The information-theoretic approach, Shannon (1948)	179
7.1	A four-category ontology, from Lowe (2006a)	202

List of Tables

1.1	Subject-Predicate structure of copular sentences (Higgins 1979, 264)	15
1.2	Composition of Subject and Predicate in copular sentences (Higgins 1979, 264)	16
1.3	Four sentence types with the same flankers	16
1.4	Roy’s taxonomy of predicational copular sentences, Roy (2013)	22
1.5	<i>The winner of the election might have been the LOser</i> —eight readings (Higgins 1979, 271–3)	31
5.1	Distribution of auxiliaries and predicate form and interpretation in Spanish, Roy (2013, ch.6)	165
6.1	Functions of copular sentences and flankers, derived from Higgins (1979) . . .	175
8.1	‘Specificational copular clauses’ Hartmann (forthcoming)	236
8.2	‘Predicative copular clauses’ Hartmann (forthcoming)	236
8.3	Three kinds of copular sentence	245
8.4	<i>Brian, the culprit</i> : possible permutations	249
8.5	<i>What Brian doesn’t eat</i> (= ‘Brian’s eating dislikes’, ED), <i>food for the dog</i> (= ‘commercial dog food’, DF): possible permutations	251
8.6	<i>The winner of the election might have been the LOser</i> —eight readings (Higgins 1979, 271–3)	253
8.7	<i>The winner of the election might have been the LOser</i> : possible permutations	255

Chapter 1

Introduction: the taxonomy of copular sentences

1.1 Aim of this thesis

The initial problem that gave rise to this thesis was how to explain the taxonomy of copular sentences.¹ Linguists generally take Higgins (1979) as an initial reference point, though there are antecedents in the work of Halliday (1967), Akmajian (1970) and also, as Higgins says, in the work of philosophers ‘whose concern with it somewhat antedates that of linguists’ (208), notably Geach (1968) and Kripke (1972). But, in line with the linguistics literature, I’ll take Higgins’s taxonomy as a reference point in this thesis too:

- (1) Predicational: *Brian is a clever guy.*
Specificational: *The culprit is Brian / Brian is the culprit.*
Identificational: *This is Mary, That man over there is Brian / Brian is that man over there.*
Equative/Identity statement: *Cicero is Tully / Tully is Cicero.*
(examples from den Dikken (2006a))

My initial premise was that it was in need of deeper explanation, and particularly that logical representations could achieve this (a solution resorted to by much of the semantic and syntactic literature). As I explored this option, I came to realise that the taxonomy is

¹Some writers prefer to talk about the ‘copular typology’, but I will use ‘taxonomy’ as (1) it is the word Higgins uses himself, (2) ‘typology’ has a specific meaning in linguistics which relates to comparisons of *languages* i.e. so that we can say what sort of a language a given language is in relation to some structural or functional property. However, the comparisons here are of different kinds of *sentence*, so that we can say what sort of a sentence a given copular sentence is in relation to other copular sentences (in the same and different languages).

really a taxonomy of *uses* of language, and as such we can't explain it, more so than *describe* it in more or less accurate ways (more on this in chapter 6). So then the problem became how to produce a better description of possible copular sentences. That, then, is the question which this thesis ultimately aims to answer: Does the taxonomy in (1) provide an adequate description of the different possible meanings of copular sentences, and if not how can it be improved? Answering this question forces us to think about what the taxonomy is based on, what factors motivate the distinct meanings we see in (1) (e.g. what makes a sentence specificational rather than predicational).

The structure of the thesis reflects this process of discovery, more on which at the end of this introduction, 1.4.

1.2 The basic data

So this thesis is an attempt to understand what categories of copular sentence there are and what these categories are based on. This may turn out to be more elaborate or less elaborate than Higgins's original four categories; of course Higgins may have hit the nail on the head and got it exactly right first time, though the literature following his work has generally sought to reduce it in some way e.g. Moro (1997), den Dikken (2006a).

A 'copular sentence' is a term from descriptive grammar that designates a particular construction in languages that use copulas (such as the word *be* in English). Copular sentences are amongst the simplest standalone propositions in English, essentially of the form:

(2) X is Y .

I'll refer to the constituents that go in the positions of X and Y as 'flankers' of the copula; it isn't a very usual term in the literature but I've found it a useful theory-neutral term. Copular sentences are by no means cross-linguistically universal. As we'll see in section 4.2, it is very common for languages to use sentences without copulas, 'nominal sentences':

(3) X Y .

In fact, English doesn't always need a copula either: we can say *I consider Brian a clever guy* instead of *I consider Brian to be a clever guy* (see chapter 5 for more details about this). It should also be noted that there are other constructions in English that can be categorised in the same way as a copular sentence (i.e. as predicational, specificational etc.). For example, structures with *as*:

(4) I am speaking to you as a friend.

Table 1.1: Subject-Predicate structure of copular sentences (Higgins 1979, 264)

Type	Subject	Predicate
Identificational	Referential	Identificational
Identity	Referential	Referential
Predicational	Referential	Predicational
Specificational	Superscriptional	Specificational

In English, (4) can also be considered as containing a predicational constituent *you as a friend*, on a par with *You are a friend* (see den Dikken (2006a) for a syntactic analysis of such constructions).

Higgins’s taxonomy was not the first and is not the only one, but it has become something of a reference for linguists and so it is convenient to take it as a starting point; one recent extension I’ll look at below is Roy (2013). A notable absence from the taxonomy are sentences with prepositional phrases like *John is in the garden*. Higgins (1979, 288) readily admits this in a footnote where he gives some clues as to how PPs might be included: ‘The locative sentences are perhaps of a separate type, or a sub-class of the Predicational type’, a comment that proved to be prescient since Roy (2013) does indeed classify such sentences as a sub-type of predicational, which she terms ‘situation descriptions’. But the point of this thesis, in any case, is not to take Higgins’s or anyone else’s taxonomy as a fact of the matter; in trying to understand the taxonomy more deeply, it may well turn out that a more elaborate taxonomy is a better description of copular sentence type variation. Higgins’s taxonomy, then, is just a useful starting point.

Higgins classifies his taxonomy according to what he calls ‘functions’ of the flankers, see Table 1.1. To be a ‘predicational’ copular sentence for Higgins is just to be a copular sentence whose subject is ‘referential’ and whose predicate is ‘predicational’. ‘Subject’ and ‘predicate’ are strictly linguistic notions for Higgins. A function is not a well-defined notion in Higgins; in chapter 6 I spend time trying to understand this notion better. Higgins nowhere formally defines these functional terms, though as I go through the thesis I’ll try to shed some light on them, particularly the notion of ‘referential’ he takes from Wiggins (1965) and Geach (1980a) in section 3.3.5.1.

Higgins gives a further analysis of what word classes can fulfil the flanker functions, see Table 1.2.² Where Higgins is clear that the sentence types of (1) reduce to statements of flanker composition, it is unclear what the relationship is between the flankers and the word classes. Flankers cannot be reduced to word classes because they can, for example, all be

²NP = noun phrase, AP = adjective phrase, VP = verb phrase, S = sentence.

Table 1.2: Composition of Subject and Predicate in copular sentences (Higgins 1979, 264)

Constituent Type	Referential	Superscriptional	Predicational	Specificational	Identificational
Deictic	+	–	–	+	+
Proper name	+	–	–	+	+
Definite NP	+	+	+	+	+
Indefinite NP	?–	?–	+	+	+
AP			+	+	–
VP,S			–	+	–

Table 1.3: Four sentence types with the same flankers

Subject	Predicate	Sentence type
definite NP (referential)	definite NP (identificational)	= identificational
definite NP (referential)	definite NP (referential)	= identity
definite NP (referential)	definite NP (predicational)	= predicational
definite NP (superscriptional)	definite NP (specificational)	= specificational

fulfilled by definite descriptions, as shown in Table 1.3 and the following examples (on the most natural interpretations):³

- (5) a. The evening star is the morning star.—identity
 b. The culprit was the guy we saw coming out of the window.—specificational
 c. The Queen is also the head of the church of England.—predicational

1.2.1 Debates about right categorisation

1.2.1.1 The debate about specificational copular sentences

The inspiration for chapters 2 and 3 is the debate within Universal Grammar surrounding Higgins’s specificationals deriving largely from the work of Moro (1997) and Heycock and Kroch (1999).⁴ Moro proposed that specificationals like

- (6) The culprit is Brian.

³Whilst Higgins allows for identificational to be constructed out of two definite descriptions according to his table, he doesn’t give an example, and it isn’t obvious what an example of this would be.

⁴What I’m calling ‘Universal Grammar’ is usually referred to as ‘generative grammar’, though Chomsky has recently clarified that ‘generative grammars’ are descriptions of individual languages, like English or Swahili; Universal Grammar (UG) is the framework within which generative grammars are developed: ‘The theory of the genetically-based language faculty is called Universal Grammar; the theory of each individual language is called its Generative Grammar’ (Chomsky, 2016).

(9) Your opinion of Edinburgh is my opinion of Philadelphia.

Under *consider* this is generally held to be unacceptable unless the copula is present:⁸

(10) I consider your opinion of Edinburgh *(to be) my opinion of Philadelphia.

Specificationals are generally regarded as being similar to equatives in this respect (examples from den Dikken (2006a, 1)).⁹

(11) Imogen considers the best candidate *(to be) Brian.

And this is unlike predicationals which can happily appear with no copula:

(12) Imogen considers Brian (to be) the best candidate.

Arregi et al. (2018) claim that the debate is centrally between the inverse predication theory, i.e. the view that the specificational subject is semantically a predicate, and the theory that the specificational subject stands for an ‘individual concept’ in Romero (2005). They consider the ‘equative’ view of specificational sentences to be a third position. This is because they define the equative view in terms of both flankers being $\langle e \rangle$ -type.¹⁰ However this is not the use here and from what I can determine authors seem quite unanimously to allow for true equations between $\langle e \rangle$ -type expressions and non- $\langle e \rangle$ -type expressions. Indeed, this is precisely what I take authors like Romero (2005) and Comorovski (2007) to be claiming: that ‘specificational’ can be reduced to *equation* between $\langle e \rangle$ -type expressions and ‘intensional objects’ (standing for the specificational subject); this is different from equations between two $\langle e \rangle$ -type objects. So what I’m claiming to be an ‘equative’ analysis is simply an analysis that involves the logical primitive of equation (‘=’) as the basic formulation of an apparent relation between the two flankers of a copular sentence. This is distinct from a predication analysis, which involves no such equation. In this sense, therefore, Arregi et al. (2018),

⁸A word about the acceptability judgement symbols in examples. Linguistics convention is that an asterisk *inside* brackets means that the sentence is unacceptable for syntactic reasons if the bracketed word or words are included; where the asterisk *precedes* the bracketed word or words (as it does in (10)), it means the sentence is unacceptable if the bracketed word or words is *not* included. Where a constituent is simply in brackets without any asterisks inside or outside the brackets, the sentence is acceptable for syntactic reasons with or without the constituent. A pound sign ‘#’ before a sentence indicates semantic/pragmatic ill-formedness, as opposed to the ill-formedness due to syntactic reasons indicated by asterisks. Question marks are used to indicate that the acceptability is unclear (if the context is unclear, I’ll clarify if this unclarity is due to syntactic or semantic/pragmatic reasons).

⁹Readers may not find the unacceptability judgement particularly sharp here. The clearest examples of this sort that I’ve found in the literature involve the use of *seem*:

(i) Brian seems (to be) the best candidate.

(ii) The best candidate seems *(to be) Brian.

(examples adapted from den Dikken and O’Neill (2017, 16))

¹⁰See the end of section 2.2.4 for the notion of simple semantic types.

following Romero (2005), is also an equative analysis, as their ultimate formulation of the meaning of a specificational sentence is a logical equation: *The lawyer is Clara* translates into ‘the individual lawyer concept (the unique x where x is the concept of a lawyer in the world of evaluation) = Clara’.

An actual argument against Arregi et al.’s way of categorising their position (i.e. excluding their own analysis from the ‘equative’ analyses) comes from the ambiguous example from Higgins (1979) that we’ll look at shortly in section 1.3. There Higgins points out that there are sentences with attributive subjects and referential predicates (based on the referential–attributive distinction of Donnellan (1966), see below 1.3 and section 8.2.2) that can be equative. And they are ambiguous with a specificational reading for the same sentence where the predicate is referential:

- (13) *The winner of the election might have been the loser.*
- a. ‘Whoever won the election might have been the same person as McGovern’—
equative reading
 - b. The one who might have won the election was McGovern.
= The following person might have won the election: McGovern.—specificational
reading

Whilst Arregi et al. don’t resort to the notion of attributive reference (or the notion of superscriptional subject that enters into Higgins’s specificational sentences), I still think they would need to make this distinction in their own idiom. In other words, they would need to show how their specificational sentences with individual concept subjects and referential predicates are distinct from equatives with the same flanker specifications, given that this reading appears to be available. This could potentially be taken to be a rather general argument against treating specificationals as identities. After all, if they were, then why is there the ambiguity in (13)? In chapter 3 we’ll conclude that only (13a) could be called an identity statement consistent with any notion of an identity statement from the philosophical literature.¹¹

1.2.1.2 Are identificationals predicational or not?

Identificationals, like we saw with specificationals in the previous section, have also been subject to arguments about whether they can be considered as predications or identities.

¹¹A reader might immediately reply that (13a) cannot be an identity statement because the modal *might* (understood with the ‘if things had been different’ reading that Higgins stipulates) implies that the flankers are non-identical. I’ll simply follow a standard linguistic line about such examples and say that there is an identity sentence at the core of (13a) which is then embedded in the scope of the modal verb *might*.

Heller and Wolter (2008) show that ‘simple’ identificationals (*That is Rosa* as opposed to *That woman is Rosa*) should actually be classed amongst predicationals. They set their evidence up as a contrast with specificationals (and the assumption that specificationals are *not* predicationals):¹²

- (14) a. Rosa is a doctor and Matilda is too.
 b. *My next-door neighbor is Rosa and your next-door neighbor is too.

(14a) shows that in a typical predicational, deletion of the post-copular phrase in a coordinate structure is licit, whereas in a typical specificational, it is illicit (14b). Identificational pattern with predicational here; the deletion is licit:

- (15) a. (pointing at pictures) That is Rosa and that is too.
 b. That is a woman and that is too.

Another argument Heller and Wolter give is that identificationals (18) also seem to pattern more closely with predicationals, (16) below, with respect to extraction from their post-copular constituent; it is relatively acceptable with them whereas it is more clearly ungrammatical with specificationals (17):¹³

- (16) a. John said that what Mary was looking at appeared to be a picture of a cat.—
 predicational
 b. ?What did John say that what Mary was looking at appeared to be a picture
 of?—relatively acceptable extraction from predicational
- (17) a. They said that what Mary was going to do was give the dog to John.—specificational
 b. *Who did they say that what Mary was going to do was give the dog to?—
 unacceptable extraction
- (18) a. Rosa said that that was a friend of John.—identificational
 b. Who did Rosa say that that was a friend of?—acceptable extraction

On the other side of this debate, Beyssade (2009) notes that in French the demonstrative

¹²It should be pointed out that the judgement Heller and Wolter give for (14b) is dubious. To me it seems perfectly well-formed and Beyssade and Dobrovie-Sorin (2009, 7–8) say: ‘The problem is that Heller and Wolter’s grammaticality judgments... are questionable... According to Louise McNally (p.c., December 2008), the examples in [(14b)]... are acceptable, which means that specificationals behave on a par with predicational with respect to the deletion of the postcopula phrase’. So it isn’t that clear that one can in fact use this data to say that identificational are somehow distinct from non-predicational sentences with respect to ellipsis.

¹³I am taking Heller and Wolter’s word for it that the sentences they tag predicational and specificational are so, but they certainly aren’t the most obvious examples of them.

ce ‘that’ is bad in left dislocation structures with clearly predicational predicates (adjectives and bare singulars in French):

- (19) a. *Marie, c’est belle. a.’ Marie, elle est belle. (predicational)
 Marie, **ce*/she’s beautiful
 b. *Marie, c’est professeur. b.’ Marie, elle est professeur. (predicational)
 Marie, **ce*/she’s a teacher

On the other hand, *ce* is good with other predicates. This includes identificational ones (20a), but also equatives and specificational. So identificational do not align with predicational on this evidence. For Beyssade, identificational, equatives, and specificational form a single category in contrast to the predicational:

- (20) a. Ca, c’est John. a.’ *Ca, il est John. (identificational)
 That, *ce*/*he is John
 b. Clark Kent, c’est Superman. b.’ ??Clark Kent, il est Superman. (equative)
 Clark Kent, *ce*/??he is Superman
 c. Le probleme, c’est John. c.’ *Le probleme, il est John. (specificational)
 The problem, *ce*/*it is John

Again, this gives a taste of the sort of linguistic evidence that has been brought in to try and refine the categorisation of Higgins’s taxonomy. These debates have persisted for a number of years with no clear resolution about which category to bracket specificational and identificational with:¹⁴ predicational (understood ultimately as having the semantics of logical predication) or equatives (understood as having the semantics of logical identity). Chapters 2 and 3 try to show that attempting to base the taxonomy on logical predication and identity is ill-founded in the first place, which perhaps may be why the debates have not yet been resolved.

1.2.2 Extending Higgins’s four categories to six, Roy (2013)

A notable recent extension to Higgins’s taxonomy is the work of Roy (2013), which subdivides the predicational class into three different types (‘characterisation’, ‘definition’, and ‘situation description’). Here I’ll spend some time introducing Roy’s theory because I refer back to it at several points in the rest of the thesis.

Since at least Carlson (1977), it has been widely accepted in semantics that there are two distinct ways of spatio-temporally distinguishing predicates: i(ndividual)-level and s(tage)-

¹⁴There is more consensus that inversion is involved in specificational (Heycock, 2012).

Table 1.4: Roy’s taxonomy of predicational copular sentences, Roy (2013)

characterisation	[<i>XP</i>]	dense	<i>Jean est professeur</i>
definition	[<i>ClP</i> [...]]	non-dense	<i>Jean est un professeur</i>
situation description	[<i>NumP</i> [...]]	maximal	<i>Jean est dans le jardin</i>

level predicates, and Roy is responding to this distinction with her three-way analysis. A ‘temporal stage of an individual’—originally due to Quine (1960)—is given by Carlson as:

A stage is conceived of as being, roughly, a spatially and temporally bounded manifestation of something. . . An individual, then, is (at least) that whatever-it-is that ties a series of stages together to make them stages of the same thing. (Carlson, 1977, 115)

Carlson gives the example

(21) Bill ran

noting that it is ambiguous between, ‘. . . an occasion on which Bill engaged in a certain activity. . . a “happening” . . .’ (the s-level interpretation) or ‘. . . that Bill had some disposition or characteristic—that he was one who runs (habitually)’ (the i-level interpretation) (Carlson, 1977, 118).

Roy’s classification of predicates, then, is a three-way rather than a two-way one, see Table 1.4, based largely on an analysis of French (but also extended to Spanish and Russian). From the table we can see that there are two ways of describing the distinctions: semantically (dense vs. non-dense vs. maximal predicates) and syntactically (*XP* vs. *ClP* vs. *NumP*). This reflects ‘a very tight connection between semantics and syntax’ (Roy, 2013, 91).

1.2.2.1 A distinction between bare noun predicates and predicates introduced by an indefinite article in French

The empirical data that grounds the three-way distinction in Table 1.4 involves two contrasts: between French bare noun predicates and indefinite article headed predicates, on the one hand, and between bare noun predicates and predicates of all other nonverbal categories, on the other. To start with the first contrast, French, like other Romance and Germanic languages (though, as noted above, not English) has an optional indefinite article in examples like:

- (22) a. Raymond est un acteur. French
 Raymond is an actor
 ‘Raymond is an actor.’
 b. Raymond est acteur.
 Raymond is actor
 ‘Raymond is an actor.’

We shall henceforth refer to (22a) as the bare noun (bare N) case, and (22b) as the indefinite + noun (indef+N) case. However, this isn’t a case of true optionality. (22a & b) are semantically distinct in a way that can be brought out from a number of tests. For example only (22b) can be an answer to the question *Qu’est Raymond?*, ‘What is Raymond/What does Raymond do?’

- (23) Qu’est Raymond?
 a. *Raymond est un acteur.
 b. Raymond est acteur.

In contrast, only indef+N is a possible response to the question *Qui est Raymond?*, ‘Who is Raymond?’¹⁵

- (24) Qui est Raymond?
 a. Raymond est un acteur.
 b. *Raymond est acteur.

A second difference between bare N and indef+N is that indef+N exhibits ‘lifetime effects’ in the past tense. (25a) must entail that Paul is now dead, whereas this need not be the case for (25b); Paul could simply no longer be practising medicine.

¹⁵For Higgins (1979, 244ff.), it seems that answering a *Who is X?* question is the hallmark of an ‘identificational’ sentence. These aren’t only sentences with demonstrative subjects for Higgins. His Table 1.2 allows for identificational subjects to be proper names, definite NPs, indefinite NPs, as well as deictics (i.e. demonstratives). They could all potentially fulfil the functional definition Higgins (1979, 237) gives of an identificational in terms of ‘teaching the names of people or things’, which allows for him to say that *John is a teacher* could be an identificational sentence, if it is used to teach someone what a teacher is (i.e. the ‘name’ of John is ‘a teacher’). This would be equivalent to *That is a teacher*, perhaps pointing to someone standing at a blackboard in front of a class of students, though Higgins (1979, 244) admits *John is a teacher* in reply to *Who is John?* would be a ‘somewhat reticent answer (of a kind dear to small children)’, unlike a more ‘helpful’ answer like *John is a teacher who’s been helping me with my polynomials*. Roy (Roy, 2013, §3.2.4.1) explicitly disagrees that sentences with indefinite post-copular expressions like *Raymond est un acteur*, in reply to *Qui est Raymond?*, are identificational, more so than predicational. She doesn’t provide specific arguments for this view there, but I take it that the overall success of her theory of nonverbal predication is intended to serve as the main argument for this position.

- (25) a. Paul était un médecin.—lifetime effects (Paul is now dead) French
 Paul was a doctor
 ‘Paul was a doctor.’
 b. Paul était médecin.—no lifetime effects (Paul may still be alive)
 Paul was doctor
 ‘Paul was a doctor.’

Thirdly, only the bare N predicates are compatible with ‘marked aspect’, where perfective aspect (the sort that could be marked in English by e.g. a present perfect tense verb like *John has eaten three cakes this morning*) is said to be marked here compared to the imperfective aspect in (26a):

- (26) a. Paul était (un) champion olympique.—imperfective/unmarked aspect French
 Paul was a champion olympic
 ‘Paul was an olympic champion.’
 b. Paul a été (*un) champion olympique.—perfective/marked aspect
 Paul has been a champion olympic
 ‘Paul has been an olympic champion.’

Finally, only the bare N predicates can be modified by temporal modifiers that ‘impose temporal limitations’ on the predicate, giving rise to an ‘interruptive reading’ as in:

- (27) a. Max est étudiant le jour, gardien de sécurité la nuit. French
 Max is student the day, guard of security the night
 ‘Max is a student during the day, and a security guard at night.’
 b. Max est un étudiant (*le jour), et un gardien de sécurité (*la nuit).

(27a) can mean—as the translation indicates—that the predicate can hold in a temporally delimited way (the day, the night). Inclusion of such temporal modifiers renders the indef+N predicate sentences unacceptable (27b).

This is some of the evidence from French that motivates the first distinction between the predicates with an indefinite article—which Roy terms ‘definitions’—and the bare noun predicates—the ‘characterisations’. Roy presents persuasive evidence to argue that this distinction is not the same as the i-level vs. s-level distinction (Roy, 2013, 3.3), and in any case, her analysis presents a three-way distinction, so now let’s look at the third category she identifies—situation descriptions—from a comparison with characterisations.

1.2.2.2 A distinction between characterisations (bare noun predicates) and the third category: situation descriptions

So far, then, we have characterisations (bare noun predicates in French) and definitions (with an indefinite article). The third category can be compared with the characterisations. The main test for this is the felicity of the example after questions that ask ‘What’s happening?’. Compare the two responses in (28):

- (28) Qu’est-ce qui se passe dehors, c’est quoi tout le bruit?
 ‘What’s happening outside, what’s all the noise?’
- a. Paul est ivre, il a renversé la poubelle.
 ‘Paul is drunk, he knocked over the dustbin.’
- b. #Paul est ivrogne, il a renversé la poubelle.
 ‘#Paul is a drunkard, he knocked over the dustbin.’

The adjective *ivre* in (28a) is perfectly acceptable, whereas the bare noun version of the adjective in (28b) *ivrogne* is not.

Also, the mere fact that the following is not a contradiction demonstrates that the bare noun predicate and adjective form must have different truth conditions:

- (29) Paul est ivrogne, mais là (exceptionnellement) il n’est pas ivre.
 ‘Paul is a drunkard, but now (exceptionally) he is not drunk.’

Finally, compare the adjective with the bare noun version with the locative prepositional-phrase modification in the following:

- (30) a. Paul est ivre dans la cuisine. ‘Paul is drunk in the kitchen.’
 b. #Paul est ivrogne dans la cuisine. ‘#Paul is a drunkard in the kitchen.’

Whilst the only examples we have looked at here have involved an adjective compared to a bare noun version, the parts of speech that can serve as the predicate for situation descriptions also include prepositional phrases (so *Paul est dans le jardin* ‘Paul is in the garden’ is also a situation description).

1.2.2.3 Treating stative predications as containing an event argument

A neat way to encapsulate Isabelle Roy’s system is as she puts it herself (Roy, 2013, 113): ‘Assuming that all predicates are predicates of eventualities (and therefore have an eventuality argument), maximal predicates range over maximal eventualities, non-dense predicates over atomic ones, and dense predicates over (unstructured) mass eventualities’. The start-

ing point for unpacking what that means is therefore to justify the assumption that event arguments are present in copular sentences. ‘Event’ and ‘eventuality’ will essentially be interchangeable terms here. Roy adopts the locution (following Bach (1986)) ‘eventuality’ as an umbrella term for both event arguments and state arguments, which she considers ‘variants of the same object’—the same ‘underlying Davidsonian argument’ (Roy, 2013, 18 n.9). Event arguments have since Davidson become a standard part of the formal semantics toolkit.¹⁶ However, postulating them for stative predications (as nonverbal predications—such as copular sentences—are) is not immediately obvious. After all, Davidson originally introduced event arguments to deal with action sentences, and ‘neo-davidsonian’ representations explicitly resort to thematic roles (for theta roles, see section 2.3.1). Theta roles are part of the lexical semantics of verbs, but as we will see in chapter 4, the status of the copula as any kind of verb in the true sense is dubious, and it doesn’t seem to subcategorize for theta roles. I shan’t go through the arguments Roy herself presents in defense of the presence of a Davidsonian event argument in these sentences (see Roy (2013, 2.4)); for the sake of presenting the rest of her system, let’s accept the conclusion that this is a licit way of representing the meanings of these sentences.

1.2.2.4 Definition as maximal predications

‘Defining’ predications are defined semantically by Roy (2013, 45) as ‘maximal’ predications, understood as referring to maximal *eventualities*. $P(e)$ is a maximal predication iff:

$$(31) \quad \forall e P(e), \text{ there is no } e' \text{ such that } e \text{ is a proper part of } e' \text{ and } P(e')$$

This essentially means that in ‘defining’ predications, reference is to the largest eventuality where the predicate P holds. To say that *Paul is an actor*, qua a definition (the equivalent of the French *Paul est un acteur* with the indefinite article), is to say that ‘Paul is in the maximal eventuality of being an actor’. To break the sentence down in a neo-davidsonian way requires introducing a ‘Max’ operator to take the place of the existential quantifier, so that $\text{Max}(e)$ essentially means ‘there exists a maximal eventuality’ (this is adapted from von Stechow (1996)):

$$(32) \quad \begin{aligned} &\text{Paul est un acteur} \\ &= \text{max}(e) \text{ Actor } (e) \ \& \ \text{Subj}(e, \text{paul}) \\ &= \text{‘Paul is in a maximal eventuality of being an actor.’} \end{aligned}$$

¹⁶See for example the essays collected in Davidson (2001); the tradition of treating verbs as denoting actions by analogy to the way nouns denote things stretches back to the ancient Indian grammarian Pāṇini (4th century BC) as noted by Parsons (1990).

‘...meaning that there is no bigger eventuality such that Paul is also in an eventuality of being an actor in that eventuality as well’ (Roy, 2013). The basic idea, then, is that the ‘being an actor’ eventuality is the ‘biggest’ in some sense that it can be, it is a maximal acting eventuality. Whilst Roy does not use the notion of a temporal ‘interval’ in (31), she does later say, when referring to the ungrammaticality of maximal predicates with temporal restrictions—such as the adverbials in (27) that express temporal limits that we saw earlier—: ‘They [maximal predicates] are... not compatible with any temporal distinctions that restrict the predicate to smaller intervals within the maximal interval I for which P is true... [A] defining predicate refers to the maximal eventuality of being P, and it is by definition not possible to restrict it to smaller intervals. Maximal eventualities cannot be relativized to intervals in time smaller than the largest interval’.

From this then I am assuming that Roy’s definition of ‘maximal eventuality’ can be broken down into something like ‘the maximal interval where the predicate holds of the eventuality’. I say this for a couple of reasons: firstly, the notion of an interval figures prominently in the definition of the other two types of copular sentence we’ll see just below, and it is otherwise puzzling that it should feature in these definitions but not be included in the maximality definition. If we are to see how these three classes of sentence are distinct from each other, I think it is important to control for all the elements that compose their definitions. The text seems to indicate that intervals are indeed relevant to the notion of maximality, as just noted in the last paragraph. The second reason to want to include the notion of interval in the definition of ‘maximal eventuality’ is simply that it isn’t obvious what a maximal eventuality is if it isn’t an eventuality of something within a maximal interval. In fact, strictly speaking, the predicate of the eventuality, say ‘the eventuality of being a student’—to take (27)—would hold ‘completely’ within this maximal interval (i.e. from the beginning to the end of it without gaps). The question that is then immediately raised is ‘What is this maximal interval?’/‘How does a speaker/hearer understand what the maximal interval is?’. In a sense, one could answer here that it doesn’t matter, so long as one understands that the interval is maximal. However, I will postpone discussion of the identification of intervals until we encounter them again in connection with the notion of density in the next section, 1.2.2.5.

1.2.2.5 Density—characterisations = non-dense predicates; situation descriptions = dense predicates

The other two categories (characterisation and situation descriptions) are understood as differing with respect to a feature that Roy calls ‘density’. Situation descriptions are dense predicates, where characterisations are non-dense. Let’s start by looking at the definition of

a dense predicate (Roy, 2013, 75):

- (33) If a predicate P is interpreted as *dense*, then P is true of an eventuality e in an interval I if and only if for any I' , a subinterval of I , there exists another eventuality e' such that P is true of e' and e' is part of e .

As already noted for maximal predicates, Roy herself never actually says how we are to identify the intervals she resorts to in her definitions of the semantics of the three predicate types (maximal, non-dense, and dense). Occasionally she refers to the ‘relevant’ interval. Do we have intuitions about what the relevant intervals may be for, say, the characteristic interpretation of a sentence such as ‘John is a teacher’ in a context where, say, Peter and Mary are discussing what John does for a living? It seems that the predicate itself dictates what the relevant interval for it would be. So, in the case of *teaching* characteristically, we would normally consider that it starts at the time John maybe completed a teaching qualification and started his first teaching job. There would be no endpoint in the present tense. In the past tense the endpoint may be the point when John retired, or when he changed professions and switched to doing another job. For other characteristic predicates, like ‘John is intelligent’, which pattern with i-level predicates, it seems like the only obvious interval is something like ‘all the time’.

The definition of a dense predicate given above in (33) can be paraphrased as meaning that the predicate is true of an eventuality (i.e. is true; all the predicates in Roy’s system are predicated of eventuality arguments) only if the ‘evidence’ that makes the predicate true holds for the entire duration of the eventuality without interruptions. What does Roy mean by ‘evidence’? Rather like with the notion of an interval, Roy nowhere formally explicates this notion. However, she does mention at a few points in the book, that the distinction between dense and non-dense predicates on the one hand, and maximal ones on the other, relates to the presence or absence, respectively, of ‘*perceptible* sub-parts’. Therefore, I think that we can say that at a minimum, the evidence making a predicate true should be perceptible evidence. For a predicate like *teaching*, it seems that we would run into difficulties if we maintained that teaching was something that could be described in purely perceptual terms; there’s surely a dimension of the ‘social function’ of perceived actions that is crucial to the definition of teaching. However, it perhaps can be said that some perceptible actions are necessary. The crucial feature is the presence or absence of subparts. It’s these that are absent in maximal predications, and which are present in the dense and non-dense ones.

For a typical dense predicate then, for it to be true that, say, ‘John is sick’, John must remain sick for the entire duration of the pragmatically relevant interval of the eventuality

e. If John recovers from his cold, say, at any point, then the sentence ‘John is sick’ is no longer true. As Roy says, ‘If there exists any (sub)interval where the predicate does not hold, the entire sentence is false’ (75). Essentially, where dense predicates are concerned, every sub-part must be identical, must be a sub-part where the predicate holds; there can be no sub-parts where the predicate may not be holding.

Non-dense predicates are, unsurprisingly, the opposite of dense predicates. They precisely do allow for gaps, ‘...i.e., episodes where the “evidence” that makes the predicate true does not necessarily need to hold’ (Roy, 2013, 77). *Jean est acteur* in French can be true without all subparts of the relevant eventuality having to provide evidence for actual acting. In other words, John doesn’t have to actually be acting constantly during the period of the eventuality of which acting is predicated, for it to be true that John is an actor, in the non-dense characterisation sense. This equally applies to the usual sense of describing someone as ‘intelligent’, for example. The dense interpretation of ‘intelligent’ would be rather unusual; in fact, it’s hard to see how we could say it. If one was asked the question ‘What’s happening’, it would be strange to say, ‘John is intelligent’ as a response (‘John is being intelligent/acting intelligently’ plausibly could be uttered, indicating that these locutions are more like dense predications). What *John is intelligent* means is something more like: ‘it is true to say of John that he is intelligent, even if at certain times he is not displaying evidence of intelligence’. The point is that he does display evidence of intelligence a qualifying amount of the time. What constitutes a qualifying amount of evidence? On this point, Roy is more explicit; she says that this ‘... has to be determined by world knowledge, pragmatic, and sociological factors for each predicate individually’. She gives the example of being an astronaut, where perhaps a single flight into space may constitute sufficient evidence, and similarly if one were to describe someone as an olympic champion. But for *Paul is a smoker* and *Paul is a drunkard*, we would require repetitive smoking or drinking over a ‘considerable period of time’ (77).

For the definition of a non-dense predicate, then, Roy says (77):

- (34) If a predicate P is interpreted as non-dense, then P is true of an eventuality e in an interval I if and only if:
- a. there exists a predicate P' true of e in a qualifying amount of nonoverlapping subintervals I' of I ; and
 - b. for every P , P' is the corresponding dense predicate.

The essential difference between density and non-density can therefore be put as:

- (35) For an interpretation to be true of an eventuality e in a relevant interval I , the state

that constitutes the evidence for the predicate

a.) needs to hold for all subintervals of I (dense interpretation).

b.) does not need to hold for all subintervals of I (non-dense interpretation)

1.2.2.6 Correlations with syntax

I shan't go through the way that these semantic definitions account for the linguistic data we saw earlier in sections 1.2.2.1 and 1.2.2.2, see Roy (2013) for details. The last thing I shall say about Roy's system is that she correlates these semantic distinctions with syntactic phrase categories. I don't think it's unfair to view Roy as ultimately attempting to *explain* her three predicational categories in terms of syntactic structure, but she doesn't actually explicitly make this claim herself beyond remarking upon the 'very tight connection' (91) between semantics and syntax here. The closest she seems to come to claiming explanatory priority for syntax is where she says (93): 'The variation in the internal structure of characterizing and defining predicates is the source of their interpretational differences, supporting a close relationship between structure and meaning'. This is consistent with the 'neo-constructionist' approach to lexical items which she adopts (Borer, 2005a; Borer, 2005b; Borer, 2013), whereby meaningful properties of lexical items 'arise' from the structures (e.g. the nominal structure of the predicates here) that they are inserted into, rather than being individually specified for each lexical item in the lexicon (Roy, 2013, 91).

In a nutshell then, what Roy says happens is this: an XP is initially classified.¹⁷ This means that a classifier head is merged to the XP, which projects in the normal way and forms a Classifier Phrase (CIP). Classifier heads introduce atomicity into the phrase, making it atomic (non-dense). Without the addition of the classifier head, a phrase is non-atomic/dense. Merging a Number head (Num) to an XP (classified or not) introduces maximality; it introduces a quantity of an eventuality necessarily interpreted as a maximal quantity. In other words, CIP is the locus of the non-dense predication and NumP the maximal predication. XP is the dense predication; being dense is therefore the default state of nonverbal predicates.

Crucially none of these predicates are DPs, which is what arguments (of verbs) must be.¹⁸ As Roy says (92), 'argumental noun phrases are DPs and only DPs can be arguments'. There are, then, four possibilities (for nominal predicates):¹⁹

¹⁷'XP' is any phrasal category, generally excluding verbs here (as we are dealing with nonverbal predication).

¹⁸See footnote 13, chapter 2, for DPs.

¹⁹In fact, (36a) is never attested; see Roy (2013, ch.4) for details, and also for a great deal of evidence for her postulating precisely the phrasal categories she does by way of correlations with the three types of predicate she identifies (maximal, dense, non-dense).

Table 1.5: *The winner of the election might have been the LOser*—eight readings (Higgins 1979, 271–3)

Identity	1. referential + referential
	2. attributive + referential
	3. referential + attributive
	4. attributive + attributive
Specificational	5. superscriptional + attributive specificational
	6. superscriptional + referential specificational
Predicational	7. attributive + predicational
	8. referential + predicational

- (36)
- a. [NP]
 - b. [CIP [NP]]
 - c. [NumP [CIP [NP]]]
 - d. [NumP [NP]]

1.3 An eight-ways ambiguous example

Higgins was happy to concede that certain copular sentences can be many-ways ambiguous in a way that the literature rarely tries to contend with. Higgins (1979, 271) gives the following example (taken from Kripke (1980)), where capitalization indicates the ‘item which bears the nuclear tone’ (Higgins, 1979, 125):

- (37) The winner of the election might have been the LOser.

Firstly we can ignore the ambiguity of the modal verb itself, ignoring the ‘may have been (as far as I know)’ reading in favour of the ‘it could have come about if things had been different’ reading. This still leaves us with—Higgins contends—at least eight readings, as in Table 1.5.

The first move Higgins makes is to distinguish between a referential and attributive use of the definite descriptions (in line with Donnellan (1966)’s distinction). In that case, in order to fix ideas, let’s say that the election referred to is the 1972 U.S. presidential election, such that, on the referential use, *the winner of the election* refers to Nixon and *the loser* to McGovern. The attributive use is ‘whoever they were’, the speaker unaware of their names and identities, merely assuming that someone or other must have won the election, and someone else must have lost it.

Higgins claims that there are four readings where (37) is an ‘identity statement’. One of these has both subject and predicate referential, equivalent to:

(38) Nixon might have been McGovern.

Higgins then also counts as an ‘identity statement’ three other readings: attributive subject + referential predicate (‘whoever won the election might have been the same person as McGovern’), referential subject + attributive predicate (‘Nixon might have been the same person as whoever lost the election’), and attributive subject + attributive predicate (‘Whoever won the election might have been the same person as whoever lost the election’).

Higgins then distinguishes two specificational readings: (1) with an attributive specificational predicate, (2) with a referential specificational predicate. He gives the respective paraphrases:

- (39) a. The one who might have won the election was whoever was the loser.
 = The following person might have won the election: whoever lost it.
 b. The one who might have won the election was McGovern.
 = The following person might have won the election: McGovern.

It might immediately be felt that it is not obvious how, say, (39b) differs from the ‘identity statement’ Higgins identifies just above, where there is an attributive subject + referential predicate. The key is to realise that identity and specification are not the same thing (despite efforts in the literature to reduce specification to a kind of identity (Heycock & Kroch, 1999; Mikkelsen, 2005; Heller, 2005)). ‘Superscriptional’ and ‘attributive’ do not mean the same thing, something that Higgins is at pains to point out with several empirical arguments (Higgins, 1979, 268–73). Later in the thesis (chapter 8), I’ll show that the ‘superscriptional’ category essentially amounts to the notion of being attributive *plus* having a certain ‘information structure’ role (for more on information structure, see 7.6.3).

Higgins then gives two more ‘predicational’ readings (1) with an attributive subject, (2) with a referential subject, which he paraphrases, respectively, as:

- (40) a. Whoever won the election might have lost it.
 b. Nixon might have lost the election.

Again, the reader may feel that there isn’t an obvious distinction between (40b), say, and the ‘identity statement’ where there is a referential subject and an attributive predicate. I’ll aim to clarify these distinctions later in the thesis; I shall try to clarify the way that the ‘referential/attributive’ distinction is different from the ‘predicational/specificational’ distinction, which I shall understand as essentially part of information structure.

At the very least, this example should serve as suggestive evidence that it would be challenging to seek a purely syntactic account of the different readings here. We would have to claim that there happen to be eight different underlying syntactic structures for (37), none of which manifest visible difference at surface structure. It isn't impossible, but we should be careful that at the point at which we start constructing such underlying structures, we aren't simply re-describing the ambiguities rather than explaining them.

We have now looked in some detail at the basic data that motivates this thesis (Higgins's taxonomy of four types of copular sentence, extended to six in the light of Roy (2013)) and gained a taste of the way that linguists have debated how to re-categorise the taxonomy, section 1.2.1. In terms of the next two chapters, it's important to note (as we saw in 1.2.1.1) the way that linguists have sought to reduce the taxonomy to two categories of predication or identity, notions taken ultimately from predicate logic, where they are primitives. The point of this last section of the introduction has been simply to highlight the difficulty that any syntactic account will face that purports to align the different readings of the copular taxonomy with different syntactic structures. There are apparently eight distinct readings of one and the same sentence, i.e. different readings that correspond to what is superficially a unique sentence structure. I also give this example of Higgins's in order to return to it at the end of the thesis (section 8.4.6), so that we can see how my own account fares in the light of this eight-ways ambiguous example.

1.4 A guide to the rest of the thesis

A popular approach to the analysis of the taxonomy of copular sentences has been to analyse them as either logical predications or logical identities. Chapters 2 and 3, then, are dedicated to showing that logic is either inappropriate for the analysis of the copular taxonomy (in the case of predicationals) or can only apply to a more restricted subset of them than is often assumed (in the case of the copular sentences that are considered identities of some kind). Chapter 2 does predication: the basic conclusion is that predication just is not a 'meaningful' notion in predicate logic. At this point, I'm making an assumption that whatever else Higgins's copular taxonomy amounts to, it is a description of 'meaningful' distinctions in some sense (later on in the thesis, in chapters 6 and 7, I'll be more specific about what I mean by these 'meaningful distinctions'). Chapter 2 also looks at the way that contemporary syntax has resorted to functions, in particular to define predication in syntactic terms. This latter endeavour does not appear very promising though, and certainly doesn't serve as a way of demarcating one category of copular sentence from another.

Chapter 3 looks at identity. The conclusion is quite different to chapter 2's. I go through a number of different potential meanings of identity statements in the analytic philosophy tradition. In 3.2 I show that the 'pure' notion of identity (from metaphysics) doesn't obviously lend itself to the meaning of an identity statement. In 3.3, therefore, I turn to work that has more directly addressed the meaning of identity statements, notably in Frege's *Begriffsschrift* (1879) and 'On Sense and Reference' (1892b). The *Begriffsschrift* theory had problems that Frege himself recognised. The 'On Sense and Reference' view is more promising, but it only really ever applies to *real* identity statements, that is to say equative copulars, like *Hesperus is Phosphorus*. Crucially, the chapter shows that there isn't really any philosophically substantiated way of saying that a specificational copular sentence like *The culprit is Brian* is an identity (this conclusion extends to identificational too, but I don't so systematically compare the putative identity statement meanings I encounter with identificational in the chapter); the meaning of a logical identity, or an identity statement of natural language analysed as a logical identity, is only ever some kind of *real* identity statement like *Hesperus is Phosphorus*.

So by now, we've said: (1) that the categories of copular sentences cannot be discriminated in terms of logical or syntactic predication (chapter 2), (2) that they cannot be understood in terms of logical identity, except possibly for the *real* identity statements, the equative copulars. Logic doesn't contain some kind of all-purpose notion of identity that could stretch to include specificational or identificational.

So the main aim of chapters 2 and 3 is to establish that we can't exhaustively explain the copular taxonomy in terms of logical predication or identity. But in doing this we've incidentally made a start on an eventual description of the copular sentence taxonomy. Whilst predication (logical or syntactic) may not *discriminate* copular sentences, syntactic predication may still be a core component of their description. We also have a notion of identity which may be relevant for equatives. The next thing I then do is turn to the copula itself. This follows on from chapters 2 and 3 in quite a natural way. One (popular) way of cashing out the difference between predication and identity is by saying that the copula itself is the locus of an ambiguity between a copula of predication and an *is* of identity (in English). Now, chapters 2 and 3 already establish that we can't really understand copular sentences as exhaustively logical predications or identities in the first place. Chapters 4 and 5, then, are a further rebuttal of the particular view that locates a logical ambiguity in the copula by saying: the copula isn't any kind of semantically ambiguous element in any case. Chapter 4 presents a standard linguistics view of the copula in terms of its being a functional element which is no kind of 'real' lexical verb (where it is only 'real' lexical verbs which are the kind of thing that could be semantically ambiguous between 'predication' and

‘identity’ in perhaps a similar way to the way that a verb like *to bank* can be i.e. meaning ‘to form a border or edge to’ *I’m driving down a driveway banked by boulders and wildflowers* or ‘to deposit money in a bank’ *I banked the cheque then went for lunch*). Chapter 4 also makes the straight empirical point that copulas are not cross-linguistic universals; many languages do without them whilst making apparently the same distinctions as English does with copulas. A theory that claims that the copula is essential as the locus of a semantic distinction (between predication and identity) needs to be aware of all the facts about the obligatoriness, optionality, or necessary absence of copulas cross-linguistically. Chapter 5 then goes on to give the reader an idea of how the copula should actually be analysed. I include this chapter because I think that chapter 4 might leave the reader (particularly the reader who thinks the copula is a semantically ambiguous element i.e. the reader who I’m arguing against) rather wondering what the copula does do. I can’t give a fully general analysis of the copula (if there is one), but I give two case studies from English and Spanish to give an idea of the things the copula may do.

So then we get to my own proposal. Rather like chapter 5, I’m not ultimately going to give a completely worked-out proposal, but rather a case study to show the way that I think a better description of the copular taxonomy can be achieved, in chapter 8. In order to lead up to this, I need to lay some groundwork. I need to do two things in particular. Firstly, I need to go back to Higgins’s original taxonomy and try to understand what it was a taxonomy of exactly. Higgins called his categories ‘functions’, but confessed he wasn’t sure what ‘function’ meant. Chapter 6, then, tries to understand what it might mean. In order to do this, it is worthwhile trying to see if we can quite generally state the function of natural language. Chomsky has notably expressed scepticism about the possibility of stating a function of language in the sense of a biologically evolved function, especially if that function is held to be communication. We confront Chomsky’s view with Thom Scott-Phillips’s rather different view which ultimately does accord a significant place to communication (and indeed clearly states a function of communication). Overall, I don’t think that Chomsky and Scott-Phillips are necessarily in disagreement. But the main point of all of this is to lead up to the characterisation of a sentence function (of which copular sentences are obviously an instance). A sentence, I’ll argue following Scott-Phillips, is an attempt to manipulate someone else’s mental representations by way of ‘extended social navigation’. Agreeing with Chomsky, I’ll accept that it is an action, a *use* of (I-)language (by our capacity for ostensive-inferential communication). In fact, this applies to the act of referring generally, in such a way that we can isolate parts of sentences as referring in different ways. The conclusion from all this is that the copular taxonomy can only be *described*, not *explained*. This isn’t to denigrate the nature of description as opposed to explanation; it is simply to make the point that

description is relevant to categorising copular sentence types (in terms of their meanings) in a way that explanation is not. One area where I have to confess certain ‘conceptual difficulties’ (to borrow Higgins’s own phrase, used in connection with functions) is with the relationship between sentence functions and flanker functions. One matter that can be decided empirically is whether flanker functions fully determine sentence functions. They can’t, as I show in 6.5.1, evidence that Roy (2013)’s ‘characterisation’ function is implemented with syntactically distinct flankers cross-linguistically. In fact, I think the right way to go is to abandon worrying about sentence functions, in favour of only dealing with flanker functions. This isn’t a position I justify directly, but the final two chapters of the thesis elaborate it to some extent.

This, then, leads us to chapter 7 which itself leads up to what we might call a ‘flanker-based taxonomy’ i.e. a description of copular sentences that just states properties of the flankers and then derives sentences from permutations of these properties, and ignores the copula and any kind of *sui generis* sentence-level description (though I leave it as an open question whether there are such descriptions, and what the relation is between them and the flanker-level properties). The conclusions of the previous chapters are collated and some outstanding issues (regarding the status of metaphysics and syntax) are dealt with. In many ways chapter 7 is the main conclusion of this thesis. The last main chapter, chapter 8, is a case study that implements the approach to describing copular sentences that the thesis has sought to justify. There are multiple taxonomies of copular sentences, depending on what you’re interested in about them. So if you’re interested in spatio-temporal properties of the nominal flankers, you can divide them up as Roy (2013) does, for example. If you’re interested in metaphysics, you can focus on the metaphysical categories of the flankers’ extensions. If you’re interested in the information structure of the sentences, you can describe them in those terms. One could even be interested in the property of whether one of the flankers moves in front of the other in the course of a syntactic derivation (as many syntacticians have been), though the reason why it does this may well belong to the realm of information structure more so than syntax. Beyond multiple taxonomies, however, I think that there is a ‘full description’ of the copular sentence taxonomy that is essentially a multifactorial model incorporating all the properties that can enter into making meaning distinctions between copular sentences. This, then, is the ultimate statement of how I think we should describe copular sentences and which will be the approach that I put into practice in chapter 8.

So chapter 7 leads up to a list at the end, 7.7.2, which collates all the factors that I think should enter into the ‘full description’ just mentioned. There are properties of copular sentences which fall under the general rubric of ‘referential uses’: for example, there is the difference between referring ‘attributively’ or referring ‘referentially’ (à la Donnellan (1966))

and there is the difference between referring to nominals densely, non-densely, or maximally (à la Roy (2013)). There are also properties of ‘information structure’ (which is mentioned in passing at various points in the thesis, particularly in section 2.2.3, which is about how Frege identifies which constituent in a sentence is a predicate; this is a forerunner of the notion of ‘focus’ in information structure). Information structure is categorised distinctly from the ‘use’ properties of language; it isn’t completely clear to me what gives rise to it, as I say in 7.6.3.4.

Chapter 8 then, as stated, is a case study which shows how some of the pieces of the puzzle I identified in chapter 7 can be used to provide a description of copular sentences. It focuses on the minimal pair *The culprit is John* and *John is the culprit* and tries to show how we can understand their specificational and predicational readings without resorting to any difference at the level of their semantics or syntax (modulo an obvious word order difference). So its subsidiary aim is—in line with chapter 2 in particular—to show that resorting to some kind of logic-based distinction between predication and identity is not necessary. I go through a recent analysis (Patten, 2012; Patten, 2016) in 8.2 which is in line with the way I think we should approach the copular taxonomy generally and which makes the distinction between specificational and predicational sentences a matter of ‘referential use’ (the term I use in chapter 7) and information structure. There is a counter-example to this view which I present in 8.3 and which I survey two analyses of before giving my own account, 8.4 at the end of the chapter. At the end of my own account I return to Higgins’s eight-ways ambiguous example that I gave just above, 1.3, to see how my account fares in the light of it. This gives rise to complications which demonstrate the need for further work to fully understand the source of the variation possible with copular sentences.

Chapter 2

Can the logical or syntactic notions of predication help explain the taxonomy of copular sentences?

Human languages do not adopt the principles familiar in modern logic. Rather, they adhere to the classical Aristotelian conception that a sentence has a subject and a predicate, where the predicate may be complex.

(Chomsky, 1988, 54)

2.1 Introduction

The point of this chapter and the next is to try and understand whether it makes sense in the first place to resort to the logical notions of predication and identity, by way of explaining the taxonomy of copular sentences we introduced in chapter 1. For now, let's just assume that the taxonomy is a taxonomy of *meanings* in some sense. Later on in chapter 6, I'll try to get clear about what the taxonomy is a taxonomy of exactly. On this basis, then, we are searching for an explanation of the meanings of the taxonomy of copular sentences. For example, is some part of the meaning of 'specification' explained by an analysis of specificationals implicating a logical equation between the two flankers?

Some readers may not think that 'explanation' is the right word here, but rather that I should be asking whether logic can be used to *describe* the taxonomy of copular sentences. I think it's fairly uncontroversial that logical representations could be given that describe at least some of the copular sentences, and later on in this thesis that will be the strongest conclusion that I think we can come to with respect to logic and copular sentences. However,

I believe certain linguists (especially semanticists) consider that logic really can serve to *explain* the taxonomy of copular sentences. That is to say, they believe that we have the copular sentences we do because the logical notions of predication and identity enter into the meanings of natural language; without logic, then, we wouldn't have copular sentences (or, at least, not the ones we do in fact have). This is the view I want to argue against here, the view that logic is truly an explanation of the copular taxonomy. To be clear, this is a view that considers the notions of 'predication' and 'identity' to be meaningful ones, and that these meanings come from logic. They are in fact primitives of logic and together exhaust the possible basic well-formed formulas of predicate logic:

Atomic formulas: an atomic formula is either a sentential letter [a predicate of degree 0] standing alone, or a predicate letter of degree n followed by a string of n individual symbols, or a string of the form $\alpha = \beta$, where α and β are both individual symbols. (Cohen, 2008)

So whilst predication and identity can't be explained themselves, they can serve to explain the meanings of natural language sentences, understood as being somehow derived from logical representations. Again, this is the view I oppose in this chapter and the next.

In order to see how this view pervades linguistics, consider the following: perhaps the most influential way in which the taxonomy of copular sentences has been analysed is the approach which has sought to reduce Higgins's four different classes to two classes, based on an analysis in terms of logical predication or logical identity. We see this in the analysis of specificational and identificational. For example, here is a very incomplete list of analyses in chronological order:

- (1) Specificationals are predications: (Williams, 1983b; Partee, 1986a)
- Specificationals are predications: (Moro, 1997)
- Specificationals are a subtype of equative: (Heycock & Kroch, 1999)
- Specificationals use a copula of identity. Identificational with a phrasal demonstrative subject (*that man* in *That man is Joe Smith*) are equatives: (Mikkelsen, 2005)
- Specificationals are predications: (den Dikken, 2006a)
- Specificationals are equatives. Identificational are predications: (Heller, 2005; Heller & Wolter, 2008)
- All identificational are equatives: (Birner *et al.*, 2007)
- Specificationals are predications: (Geist, 2007)
- Specificationals are equatives: (Selvanathan, 2016)
- Some specificational are equatives, others are predications: (Barros, 2016)

These linguists are resorting to (intensional) predicate logic in order to account for their taxonomy. But is this right? It isn't obvious that predication—qua a logical notion—is, or has ever been analysed as, meaningful in any sense. Identity, on the other hand, is generally viewed as a meaningful notion, but it may be that the meaning associated with identity statements of predicate logic is quite restricted. So it may only apply to natural language sentences like *Hesperus is Phosphorus*, not a specificational sentence like *The culprit is Brian*. From that perspective, claiming (à la Romero (2005) or Comorovski (2007)) that specification is an equation between an object (*Brian*) and an 'individual concept' or 'intensional object' (*the culprit*) may just be an illicit equation.

What do these linguists think the meaning of predication, in particular, is? The general consensus is that it is 'property ascription' (Heycock & Kroch, 1999; Adger & Ramchand, 2003; den Dikken, 2006a; Mikkelsen, 2011), which is considered a primitive and self-explanatory. But the kind of predicate logic that linguists resort to doesn't contain properties as any kind of primitive. It contains objects, truth values, and functions. Derivatively, we can translate functions into sets, such that we have sets of objects. There is in fact a place where 'property ascription' has long been considered to be a substantive notion—metaphysics—so in chapter 7 I'll try to show how we might describe *some* of the copular sentences as property ascriptions in those terms.

Linguists seem quite happy to talk about 'equating an object (*Brian*) with an intensional object (denoted by *the culprit*)' or 'predicating the property of being the culprit of an object (*Brian*)' without really considering whether these are licit locutions. Behind this may lie the way that they deal in syntactic operations and the view that syntax determines meaning in some way. Syntacticians are rarely explicit about stating this assumption, but we see it for example in den Dikken (den Dikken, 2006b, 303–4) [my emphasis, in bold]:

The generative literature has generally tended, however, toward the one-*be* approach, ascribing the semantic variability of copular sentences to the **structures** involved, not to the semantics of *be*. . . [O]ne should proceed on the assumption that all of the semantics of copular sentences derives from (i) its major constituents and (ii) their **syntactic structure and derivation**.

Here are some more quotes to justify the view that syntacticians assume that syntactic distinctions underpin meaning distinctions:

I shall try to show in this article that the extension of X-bar theory to functional categories makes possible a theory in which both main clause and SC predication can be given a uniform structural definition and, correspondingly, a uniform semantic interpretation. (Bowers, 1993, 591)

If we also assume, given that they are semantically near-equivalent, that (1a) and (1b) have the same underlying structure...’ (Haegeman & Guéron, 1999, 288)

A (neo-)constructionist view of the lexicon is adopted, which assumes that the three-way distinction between maximal, dense, and non-dense predicates arises from properties of the structures into which the lexical items are inserted rather than properties of the lexical expressions themselves. (Roy, 2013, 91)

These quotes could potentially be in line with a range of different views from ‘weaker’ versions where syntax and semantics merely coincide non-systematically to a strong view that semantic differences are determined or explained by syntactic ones (thanks to Ulrich Reichard p.c. for clarifying this point for me).

Two syntactic operations in particular may be relevant in relation to the view that syntax underpins meaning somehow. ‘Nominalization’ can make a nominal out of verbs (*destroy* → *destruction*), adjectives (*wise* → *wisdom*), and even copular sentences (*John is in the garden.* → *John’s being in the garden*); this may give rise (along with the premise that nominals determine object-hood somehow) to the view that any nominal is an object (e.g. Martin et al. (2016)). Another apparently syntactic operation—linguistic/syntactic predication—allows us to put almost anything in the position of syntactic predicate. This, then, allows those who view syntactic predication as determining a meaning of property ascription to say that *Hesperus is Phosphorus* may mean that there is a property ascription involved somehow (as Hinzen (2016) explicitly does, see below section 2.2.7). To be clear, this is not a view I endorse here, but it may explain why there seems to be a systematic confusion about treating copular sentences as exhaustively logical predications or identities. I should probably be more circumspect about ascribing to linguists an implicit adoption of the strong idea (notably in Hinzen’s work) that syntactic complexity determines ontological distinctions. There may well be no particularly developed notion of ‘object’ and ‘property’ underlying their locutions. But I think it’s fair to say that they are assuming a determination of meaning by syntax that goes beyond the effects of mere compositionality if they believe (as in the den Dikken quote above) that syntactic structure can account for the semantic variability of copular sentences. Syntax must apparently explain predicational (definition, characterisation, situation description) vs. specificational vs. identificational vs. identity (the taxonomy we saw in chapter 1).

So I shall take predication and identity in turn, predication in this chapter and identity in the next. I shall start by looking to see what the origins of logic tell us about putative ‘meanings’ of predication and identity. I don’t actually need to be particularly discerning here; if the literature gives rise to multiple substantive notions of predication and identity,

I'm happy to take any or all of them. The point then is to see whether any of these could be usefully taken to account for the meanings of any of the copular sentences from the taxonomy. The conclusion will essentially be: there's no substantive notion of 'predication' from logic that could demarcate a class of 'predicational' copular sentences from any others, based on their meanings. Where identity is concerned (in chapter 3), there are a few different notions of identity, and they may be appropriate for an informal description of classic identity statements, like *Hesperus is Phosphorus*, but not specificationals like *The culprit is Brian*, or identificationals like *That is a tractor*. But this chapter is just about predication.

In section 2.1.1, some philosophical views of the meaning of predication are briefly given, notably Davidson's view of predication in terms of a primitive notion of truth. Section 2.2 goes into Frege's conception of a predicate and predication in some detail, in order to see if it is a meaningful category. The second main part of this chapter, section 2.3 then goes on to see how functions have entered natural language syntax. The main point of section 2.3 is to see whether there is a syntactic definition of predication, which might serve as a definition of the meaning of predication (in terms of something like 'property ascription'); this is built up to in sections 2.3.1 to 2.3.4. The answer is that any such definition has proved difficult to achieve and section 2.3.5 backs this up, pointing out that syntax has struggled to dissolve 'predication' into notions from contemporary syntactic theory in the way it has done with other taxonomic artefacts like passives, relative clauses, and other notions from traditional grammar. Furthermore, to the extent that a syntactic definition of predication could be achieved, it would not determine the meaning of a 'predicational' copular sentence in a way that would serve to distinguish it from the others (the point which also applies to logical predication). Sections 2.3.6 and 2.3.7, somewhat incidentally, introduce two other ways in which linguistics can be understood as dealing in mathematical functions of the sort Frege resorted to.

2.1.1 Some philosophical approaches to the meaning of predication

By way of giving a little history of the term 'predication' very briefly, Plato is plausibly the first to introduce the notion of subject and predicate, when he defines a sentence in the *Sophist* as, 'a synthesis of... a nominal and a verbal expression, i.e. subject and predicate' (Seuren, 1998, 15) terms which were used as part of his correspondence theory of truth. Aristotle has the term *rhêma* which is translated as 'verb' or 'predicate' and which is the 'sign of what holds... of a subject' (*De Interpretatione* 3, 16b, 7). Aristotle then introduced two further terms, one to designate the grammatical predicate (*katêgoroumenon*) and another

to designate the property in the world denoted by the grammatical predicate (*symbebekós*). But he failed to make the same grammatical–metaphysical distinction on the side of the subject, only designating the grammatical subject with the term *hypokeímenon* (translated as ‘subiectus’ in Latin, hence ‘subject’ in English). Furthermore, the Latin translation of Aristotle conflated both *katégoroumenon* and *symbebekós* as *praedicatum*, whence ‘predicate’ in English. About this, den Dikken (2006a, 8) says:

This in and of itself does not mean that we should abandon the distinction between subject and predicate, however. Terminological confusion does not jeopardize Aristotle’s notion that a declarative sentence can be teased apart into something that ascribes a certain property to something else and the entity to which this property is ascribed (or ‘that which the sentence is about’).

I put this quote deliberately as I think it nicely illustrates two things: (1) Notice that ‘property ascription’ has crept in. (2) Notice also that it is rendered as apparently synonymous with ‘that which the sentence is about’.¹ But I think this apparent synonymy needs to be justified, as *prima facie* ‘that which the sentence is about’ could be distinct from a notion of property ascription (which, in fact, is the position I take in this thesis, see especially chapter 7²).

Aristotle furthermore claimed that propositions (combinations of subject, predicate, and tense) came in only two forms: affirmation and negation. As such, as Horn (1989) points out, tense and negation are syncategorematic (they aren’t subjects or predicates themselves, but combine with subjects and predicates). Modern syntax follows Aristotle in assigning tense and negation to functional (as opposed to lexical, see section 4.3) categories. The common view in syntax of assigning predication itself to a functional category (‘Predication Phrase’, see section 2.3.4) is also Aristotelian in essence, in so far as it can be seen as a cover-all term for the two polarities of affirmation and negation (Svenonius, 1994, 2–3). This is somewhat different to the view in modern semantics, where predication is functional application, see section 2.2, and negation is ‘a one-place operation over the denotations of sentences, which has the effect of reversing the truth-value’ (Svenonius, 1994, 4).

History aside, Reichard (2013, 34) gives some quite typical examples of how philosophers think of predication as a meaningful notion. They often refer to it as a ‘mode of combination’ which expresses a relation determined by a predicate. For example:³

¹I take this to be stylistic quoting and not a direct quotation from Aristotle.

²Essentially I’ll claim that property ascription may, if anything, be a straight reflection of a metaphysical fact about the world, whereas that sentences have parts which they are about is an information structural fact about sentences (which wouldn’t obviously exist in a world without sentences, unlike the fact of a property inhering in an object).

³How do we determine which constituent (*Theaetetus* or *sits*) is the predicate? Generally, it seems

(2) Socrates sits.

So what does *sits* mean? Here we are given a few options: *sits* could be a *universal*. In this case, predication would express the relation of ‘instantiation’. In (2) Socrates, ‘instantiates the universal of sitting’ (Reichard, 2013, 34). But we are also given two alternatives: ‘property bearing’ (Socrates may bear the property of being seated), and ‘set membership’ (‘Socrates is a member of the set of seated objects’). It is perhaps telling that we aren’t offered amongst these options the apparently most obvious meaning of predication for the example of (2), namely that Socrates is an object that performs a habitual action of sitting, or maybe is performing the action of sitting at the time of utterance of (2).⁴ One problem with this for such philosophers, presumably, would be its failure to be all-purpose enough. We can immediately see that not everything we consider a predicate will mean ‘perform a habitual action’ or ‘be in the process of performing an action at the time of utterance’, though it is more plausible that they could be productively described as universal instantiations, property ascriptions, or set memberships. It isn’t clear to me that such a generalization is necessary i.e. I don’t see why we need to generalize the meaning of (what I here consider to be) a syntactic position, the syntactic predicate position. In fact, this question (‘Should we so generalize?’) never seems to arise. I think the reason for this is that those who make these generalizations are often linguists who believe as an assumption that syntactic positions determine uniform meanings. Or they may be logicians who believe that logical predication is a meaningful notion, something I’ll try to show is, at least, not founded on Frege’s own words.

One notable strand in analytic philosophy led Davidson (2005) to argue that the meaning of a predicate should be understood in terms of *what objects it is true of*. It is not possible here to go into the details of this claim (though see Reichard (2013, ch.2) for such details). Essentially, Quine (1960) took Tarski’s (1933/1983) idea that truth can be defined in terms of the ‘satisfaction’ of predicates by objects and turned it on its head i.e. Quine said that ‘general terms’ can be understood in terms of what objects they are true of. Davidson then added that truth was a substantive notion but a primitive which could not be further defined (partly in order to avoid a circularity whereby truth may be defined in terms of satisfaction,

philosophers don’t say. So it must be a stipulation for them that *sits* is priorly determined to be the predicate of (2). Reichard himself is sceptical about the homogeneity of the notions of ‘subject’ and ‘predicate’ (personal communication), but in any case has a system in his (2013) which creates syntactic asymmetries that could be used to define such relations.

⁴This remark is consistent, I think, with a criticism of Davidson that Lowe (2006) makes, in which he states that ‘Theaetetus’s act of sitting, which is a particular event’ is something extra-linguistic which suffices for the truth of the sentence *Theaetetus sits*, contrary to Davidson, who does not think that there is anything extra-linguistic which makes *Theaetetus sits* true. So Lowe’s view is a correspondence theory of truth, in line with e.g. Moore (1910–11, ch.15).

which would then be defined in terms of truth etc.). This, very schematically, is the way that the meaning of a predicate can be understood in terms of what objects it is true of.

Davidson's view seems quite close to mainstream formal semantics, which takes predicates to be defined in terms of functions which can be 'translated' into set-talk, in such a way that we can say that

(3) Mary smokes.

contains the predicate *smoke*, which is the set of all and only those who smoke. If Mary is a member of this set, then (3) is true. Presumably, Davidson would put things the other way round and say that if (3) is true, then the meaning of *smoke* includes the fact that it is true of Mary. So formal semantics seems to use sets to say whether a sentence is true and Davidson, inversely, seems to use truth to say what the extension of a set is (understood as the meaning of a predicate). The point of the next section, 2.2, is to build up from Frege's work to a more thorough understanding of the formal semantic definition of a predicate. But I shall conclude there that it is not a semantically substantive notion. As such, if Davidson's view is indeed similar to this mainstream formal semantic view (or just an inverted version of it), then I don't think it is a semantically substantive view either.

2.2 Predication in the work of Frege

2.2.1 The meaning of a predicate is a concept

Now let's go into Frege's system more systematically (which will incidentally take us into the foundations of mainstream formal semantics). Doing this will help us to see if linguists' talk of 'the property of being the culprit is predicated of Brian' is a statement about the *meaning* of *Brian is the culprit*? It won't be obvious that it is, and hence that logical predication cannot serve as an explanatory primitive for discriminating the meanings of copular sentences.

Frege characterised what he himself referred to as the 'meaning' of a predicate. This was what he called a 'concept'. In Frege (1892a, 43), he says this: 'A concept (as I understand the word) is predicative' and then in a footnote: 'It is, in fact, the meaning of a grammatical predicate'. He goes on: 'On the other hand, a name of an object, a proper name, is quite incapable of being used as a grammatical predicate'. Giving the examples

- (4) a. This leaf is green.
b. This horse is a mammal.

Frege says that *is* is a copula, which he takes to be ‘a mere verbal sign of predication’, and he gives an example in German where the copula can be dropped and replaced by ‘the mere personal suffix’:

- (5) a. Dies Blatt ist grün.
 ‘This leaf is green.’
 b. Dies Blatt grünt.
 ‘This leaf is green.’ lit. This leaf greens

Frege continues (Frege, 1892a, 44): ‘We are here saying that something falls under a concept and the grammatical predicate means this concept’. He contrasts examples like (5) with those like (6) (he doesn’t give subjects, so I’ll just put a placeholder, *X*):

- (6) a. *X* is Alexander the Great.
 b. *X* is the number four.
 c. *X* is the planet Venus.

Here, Frege says, ‘... “is” is used like the “equals” sign in arithmetic, to express an equation’, which remark is also followed by a footnote: ‘I use the word “equal” and the symbol “=” in the sense “the same as,” “no other than,” “identical with.”’ (Frege, 1892a, 44).

So in a typical predicational sentence like

- (7) The morning star is a planet.

Frege says, ‘... we have a proper name, ‘the morning star,’ and a concept-word, ‘planet’; *the morning star falls under* the concept ‘planet’. Frege further remarks that an object’s falling under a concept is irreversible. If an object *o* falls under a concept *C*, it is not the same as a concept *C* falling under an object *o* (indeed it isn’t possible in Frege’s system for concepts to fall under objects).

What, then, is a concept?

2.2.2 Concepts as functions

The basic sentences of Frege’s (1893/1903) system are of the form:

- (8) $() = ()$

This sentence signifies a function that takes a pair of objects from its domain set and maps them to the ‘The True’ (in the terms of the *Grundgesetze*, Frege (1893/1903)) if the pair of objects are identical, and maps them to ‘The False’ if they are not identical, in such a way

that a sentence like

$$(9) \quad (2^2) = (4)$$

is mapped to The True.

Concepts are a class of this basic sort of identity sentence, but one where one of the arguments is a placeholder (like one of the ‘()’s of (8)), in such a way that the expression is incomplete, or ‘unsaturated’. So instead of (9), we can give a concept

$$(10) \quad ()^2 = (4)$$

which in theory is the concept ‘everything which is identical to 4 when squared’. Now, it just so happens that there are only two things that would map to the True in this case (the numbers 2 and -2). But other concepts can have a larger number of domain values. The concept ‘everything which is greater than 2’ can be given as:

$$(11) \quad () > 2$$

To retain uniformity with (10), this can be understood as elliptical for ‘() = >2’ i.e. the placeholder is identical to being greater than 2. In (11), the values for the placeholder which would be mapped to The True would be any number in the domain set which is greater than 2. If the domain set only contains 3 and 4, say, then it would be these two numbers. But if the domain set is intensionally defined as ‘the set of all natural numbers’, then the number of values of the placeholder in (11) will be infinite.

So it is important to notice two things about Frege’s concepts. 1.) They are functions which have as their range *truth-values*. 2.) They are functions where one of the arguments is unsaturated, i.e. is a placeholder.

Functions are essentially just mappings from a domain set of objects to a range set of objects. A given object in the domain of the function is the input and is linked by the function to an object in the range of the function. So, for example, given the domain of natural numbers, the squaring function $()^2$ takes a given natural number, say 2, from its domain and maps it to 4, 3 maps to 9, 4 maps to 16 etc. So note the subtle difference between ‘normal’ functions like the squaring function:

$$(12) \quad ()^2$$

—where given the domain input ‘2’ the range output will be ‘4’—and the *concept* ‘that which when squared is identical to 4’.

$$(13) \quad ()^2 = 4$$

Concepts are functions that rely on such identities, taking a pair of objects as their domain set and a truth value as their output, so that here given the domain input ‘2’, the range output will be ‘true’, because 2^2 is indeed identical to 4.

From this, then, it looks like we are defining predication in terms of functional application. Frege, incidentally was quite explicit about doing this in the *Begriffsschrift* (Frege, 1879, 7):

I believe that the replacement of the concepts subject and predicate by argument and function, respectively, will stand the test of time. It is easy to see how regarding a content as a function of an argument leads to the formation of concepts.

So, to move things to natural language sentences,

(14) John is happy.

contains a function that maps people to a truth value. This function can be given in the form of a lambda calculus formula:

(15) $\lambda x . \text{Happy}(x)$

The ‘ λx ’ just introduces a variable into the formula, the lambda essentially performing the role of taking an input and substituting it into the expression on the other side of the dot. If we wanted to be pedantic and maintain the same format as (8), we could write it as:

(16) $\lambda x . \text{Happy}(x) \ \& \ x = \text{John}$

This contains an identity sentence, and would be the function ‘That which is happy is identical to John’ (rather like (13) is the function ‘that which when squared is identical to 4’), which would map to true if the variable was substituted for John and false otherwise. So there would be a pair of objects ‘that which is happy’ and ‘John’ which would map to true if they are identical (just as ‘that which when squared is identical to 4 maps a pair of objects, x^2 and 4, to true if they are identical, i.e. if $x = 2$ or -2). Evidently, substituting an argument for the variable in this expression would give us:

(17) $\lambda x . \text{Happy}(x) \ \& \ x = \text{John} \ (\text{John})$
 $\text{Happy}(\text{John}) \ \& \ \text{John} = \text{John}$

This would therefore ultimately map a pair of objects $\{\text{John}, \text{John}\}$ to true. As we can see then, the part of the function after the conjunction is redundant when the sentence is true (just as a set containing two identical elements John is equivalent to a set containing only one John), and so we can innocently talk of this function as mapping John to true if

he is happy, and of functions mapping objects to truth values generally when they are true (rather than always talking of sentences containing functions that map pairs of objects to truth values according to whether they are identical or not).

This notion of a concept as a function from objects to truth values is at the root of contemporary formal semantics, as we'll see below. However, it is interesting to note as an aside that modern predicate logic has things the other way (which of course is not at all surprising if modern predicate logic is consistent with Frege's overall project to locate the foundation of mathematics (i.e. mathematical notions like 'function') in logic):

By contrast, in the modern predicate calculus, this last step of analyzing predication in terms of functions is not assumed; predication is seen as more fundamental than functional application. The sentence 'John is happy' is formally represented as 'Hj', where this is a basic form of predication ('the object j instantiates or exemplifies the property H'). In the modern predicate calculus, functional application is analyzable in terms of predication. . . (Zalta, 2017)

Zalta (2011) also gives arguments why predication should be considered more fundamental than set theory; as we'll see in 2.2.4, it is common in formal semantics to translate function-talk into equivalent set-talk.

2.2.3 How to determine the function for a given sentence

Given the above, the important—but perhaps rather overlooked—matter of how we identify which expression counts as the function—for a given sentence—arises. In the simplest cases like

(18) John is happy.

we could claim that two possible functions here are either 'John's being something' or 'Something's being happy':

- (19) a. John is (X) = true iff John has the property assigned by X
 b. (X) is happy = true iff X has the property of being happy

Strictly speaking 'John is X ' in (19a) would require ' X ' to be a variable of properties, rather than objects, so would require a second-order predicate logic which allows for such variables. In any case, we can take one of Frege's own examples in order to illustrate this point about identifying what the function is in a given sentence:

For Frege, the function is that part of the expression that is invariant, the argument being the replaceable or exchangeable part. Thus in a pair of sentences *Carbon dioxide is heavier than hydrogen* and *Carbon dioxide is heavier than oxygen* (Frege's examples...), we are dealing with the same function with different arguments, where *hydrogen* and *oxygen* are the arguments; the function is that part of the expression that does not change. (den Dikken, 2006a, 9)

So in this example, the function is 'Carbon dioxide is heavier than X '. Of course, the invariant part could be different:

- (20) a. Carbon dioxide is heavier than hydrogen.
 b. Helium is heavier than hydrogen.

In this case, the function would be ' X is heavier than hydrogen'. This is interesting and, I think, necessary in determining which function a given sentence contains. It seems quite clear that we can only have this notion of a replaceable part in relation to what we might call a 'comparison set', like the pair of sentences Frege gives. But if we are just dealing with a single sentence on its own and want to know its replaceable part, then how do we know what the comparison set is? We would presumably have to resort to wider discourse, and in fact it seems likely that for any given sentence it is actually possible to discover an implicit comparison set (later in chapter 7 I'll outline this idea when I talk about bringing in the notion of focus from the branch of linguistics known as 'information structure'). Den Dikken rightly observes that Frege's function doesn't relate to any syntactic or lexical constituent. As such, he decides it isn't useful in order to determine which element in the sentence is a 'predicate'. But later on, I'll argue that this method of identifying the function in terms of the 'invariant part' is simply a useful notion in its own right.

McGinn (2000, 9) makes a point which I think is related to the immediately above. His central concern there is to say that the notions of identity and distinctness are so fundamental that we can't conceive of any kind of predication without presupposing them. He makes this point by saying: 'When we predicate a property F of an object x two types of multiplicity are involved'. One is that we single out x and not some other object to predicate F of. The other is that F 'is but one of many properties' that might be predicated of x . So there is the plurality of objects that 'might or might not be F ' and the plurality of predicates/properties that might be predicated of x : 'But both these thoughts involve the idea of identity, since they bring in the idea of objects and properties that are *not* identical to x and F ', i.e. the notion of a plurality presupposes identity and 'its twin'—distinctness. About this, I think I would say that McGinn's point is a necessary truth that allows for the phenomenon he identifies, the 'two types of multiplicity' involved in any act of predication. A predication

is always, it seems, made in the context of other possible predications. This is the way in which a given utterance is informative, it ‘reduces uncertainty’ in Shannon et al. (1951)’s terms. Where an addressee’s information state at t_0 may be

(21) t_0 : Paul Pogba is a French, Belgian, or African footballer.

a speaker’s utterance of the predication

(22) Paul Pogba is not an African footballer.

at t_1 means that their addressee’s information state at t_1 becomes, effectively,

(23) t_1 : Paul Pogba is a French or Belgian ~~or African~~ footballer.

which is a less uncertain state of knowledge than t_0 in that there are only two disjuncts rather than three, and two possible states of the world is less uncertain than three possible states of the world. I think it is this phenomenon that McGinn identifies when he says that any given predication is founded upon the knowledge that the subject and predicate are not other subjects or predicates. His point could be made relative to informativeness: without a notion of identity/distinctness, we wouldn’t be able to judge any given predication as informative, because we would have no notion of other possible predications. ‘Other’ carries with it the meaning of ‘not the same’. ‘Other predications’ are predications which are not the same as a given predication. So McGinn’s point still stands. I would simply say about it that the interesting phenomenon (that certainly relies upon a notion of identity) is the informativeness of any given predication, understood as an act of uncertainty reduction. And, to bring things back to Frege, knowing how a sentence is informative—i.e. knowing what the information state is which a sentence reduces the uncertainty of—can help us determine what the function is in a given sentence. The function of (22) must be ‘Paul Pogba is X ’ (or ‘Paul Pogba is an X footballer’), not ‘ X is not an African footballer’ which would only be the function if the information state being reduced were:

(24) Romelu Lukaku, Paul Pogba, and Leroy Sané are not African footballers.

It’s worth noting, in passing, that McGinn’s point that predication presupposes identity is at least consistent with Frege’s reducing all predications ultimately to identity sentences of the form in (8).

2.2.4 Functions as sets

There is a possible further stage in our thinking of predicates as functions, as noted by Zalta (2011, 4):

Logicians and mathematicians following Frege recognized that (a) functions could in turn be analyzed as sets of ordered pairs, (b) that functions which map their arguments to a truth-value could be replaced by the set of the objects which are mapped by the function to The True...

We can imagine a world in which there are only three people: John, Mary, and Brian, and where John and Mary are happy and Brian is unhappy. In this case *happy* will designate the following set of ordered pairs

$$(25) \quad \textit{Happy} = \{\langle \textit{John}, 1 \rangle, \langle \textit{Mary}, 1 \rangle, \langle \textit{Brian}, 0 \rangle\}$$

The second (b) point Zalta makes is that we could potentially read off from this set of ordered pairs the set $\{\textit{John}, \textit{Mary}\}$ which would be the ultimate meaning of *happy* in this restricted world; it is the set of all happy people. One may wonder why $\langle \textit{Brian}, 0 \rangle$ is included at all in the set of ordered pairs in (25), and this same point could be made about *happy* understood as a function; why must Brian appear at all in the domain set of objects for that function? Why couldn't one just say that exclusion from the set implies falsehood with respect to the application of the function? In that case reading off the set of ordered pairs to a set of objects is an even more direct correlation; the set of happy people will be exactly the same as the set of objects that appear as the first member of the ordered pair, which would directly correspond with the set of objects that appear in the domain of *happy* understood as a function. But I accept there may be an important reason why functions and their associated sets of ordered pairs must just be exhaustively explicit with regard to which truth value any given domain object is mapped to.

This approach to functions (boiling them down to sets of ordered pairs) generalizes to sets of n -tuples generally. So *John loves Mary* being true means the ordered pair $\langle \textit{John}, \textit{Mary} \rangle$ is a member of a set of ordered pairs where the first member is itself an ordered pair i.e.

$$(26) \quad \textit{love} = \{\langle \langle \textit{John}, \textit{Mary} \rangle, 1 \rangle, \langle \langle \textit{Lucy}, \textit{Brian} \rangle, 1 \rangle, \langle \langle \textit{Harry}, \textit{Megan} \rangle, 0 \rangle\}$$

The whole edifice of formal semantics (e.g. Heim and Kratzer (1998)) is built up from these foundations. More complex functions can be built up from these basic building blocks, in such a way that the function corresponding to *love* is not simply a function from objects

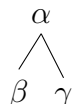
to truth values $\langle e, t \rangle$ (where ‘e’ denotes objects, and ‘t’ denotes truth values⁵), but rather a function from objects to a function from objects to truth values: $\langle e \langle e, t \rangle \rangle$.

2.2.5 Predicate modification and set intersection

For the sake of completeness, it should be mentioned that there is one other major composition rule standardly introduced in formal semantics (Heim & Kratzer, 1998), namely ‘predicate modification’:

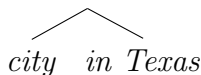
- (27) If α is a branching node, $\{\beta, \gamma\}$ is the set of α ’s daughters, and $[[\beta]]$ and $[[\gamma]]$ are both in $D_{\langle e, t \rangle}$, then $[[\alpha]] = \lambda x \in D_e . [[\beta]](x) = [[\gamma]](x) = 1$.

A ‘branching node’ just means a composed phrase like:

- (28) 

‘[[X]]’ signifies ‘the interpretation of X’, ‘ $D_{\langle e, t \rangle}$ ’ means ‘the set of all $\langle e, t \rangle$ functions’, and D_e ‘the set of individuals’ (Heim & Kratzer, 1998, 28). As mentioned before in 2.2.2, ‘ $\lambda x \in D_e \dots$ ’ in (27) just serves to introduce a variable into the formula, the lambda essentially performing the role of taking an input (in this case an $\langle e \rangle$ -type expression) and substituting it into the expression on the other side of the dot.

Applied to *city in Texas*, we get:

- (29) 

- (30) $[[\text{city in Texas}]]$
 $= \lambda x \in D_e . [[\text{city}]](x) = [[\text{in Texas}]](x) = 1$
 $= \lambda x \in D_e . x \text{ is a city and } x \text{ is in Texas.}$

This gives the meaning of a ‘complex’ predication where two predicates are applied to one object. It’s quite significant that Heim and Kratzer (1998) feel it necessary to include this new composition rule. The formal semantics programme is based on an extension of what Heim and Kratzer call ‘Frege’s Conjecture’, the idea that semantic composition is always functional application where one of the composed elements is a function and the other its

⁵Such ‘types’ were originally introduced by Russell to solve problems in the foundations of mathematics (notably a paradox he had himself identified in Frege’s set theory, frequently referred to as ‘Russell’s Paradox’) in works like his (1908).

argument. In the case of *city in Texas*, we would expect either *city* or *in Texas* to be the function that the other applies to. But this is not the case, so predicate modification is a departure from Frege's Conjecture.

Where a 'meaning' of functional application is already doubtful, it seems clear that predicate modification just serves to conjoin two predicates to one object, and so is as meaningful as conjunction, or as set intersection: *city in Texas* is the object which is both a member of the set of cities and a member of the set of things in Texas. Den Dikken (2006a) claims that the semantics of predication always involves set intersection. He summarizes his overall view (17): '[I]t will be sufficient to think of the semantics of predication as an intersective relationship between two sets, one (corresponding to the function) denoting a property ascribed to the other (the argument)'. But does it really make sense to combine set intersection and property ascription in this way? For example he says (23): 'For transitive sentences like *Imogen kissed Brian* we have assumed... a semantic representation in terms of the intersection of two sets, one denoted by the subject (*Imogen*) and one by the predicate (*kissed Brian*)'. But set intersection gives us an object, not a truth-value or any kind of sentential meaning. Here, the intersection of *Imogen* and *kissed Brian* would just be Imogen (i.e. the object that is both a thing that is Imogen and a thing that kissed Brian).

2.2.6 Intensionality

For even more completeness, one further point needs to be made about the formal semantic understanding of predication based on Frege's functions. Chierchia (1984) outlines the way that 'Montague semantics' (what is now more commonly just referred to as 'formal semantics'), 'provides a general theory of properties that relies on two major subcomponents: a possible world analysis of intensionality and the theory of simple types'. We've already seen (at the end of section 2.2.4) the theory of simple types in the basic Fregean concept function from entities (type $\langle e \rangle$) to truth values (type $\langle t \rangle$). And we saw that these building blocks could build up from the basic concept function $\langle e, t \rangle$ to more complex functions, such as the one that could be used to represent the meaning of a transitive verb like *love*: $\langle e \langle e, t \rangle \rangle$. The 'theory of simple types' makes it possible 'to generalize the theory of properties... from basic individuals... to everything we might possibly want to talk about (including any higher order entities) in a paradox-free way' (Chierchia, 1984, 2). So not only do we have a 1-place $\langle e, t \rangle$ property of basic entities, but we have a 'property of properties of basic entities' $\langle \langle e, t \rangle t \rangle$, and so on, the general rule being that in any $\beta(\alpha)$ predicative structure, β will have to be a higher type of property than α .⁶

⁶see e.g. Gamut (1991a, 122) for a list of intensional semantic types.

With respect to the other ‘major subcomponent’ of formal semantics—the possible world analysis of intensionality—Chierchia says, ‘a property is analysed as a function from possible worlds into extensions (sets or characteristic functions of sets)’. So far we have said nothing about intensionality; our outline of Frege’s functions was purely extensional. However, the current mainstream in formal semantics renders Frege’s functions intensionally. So, as Chierchia puts it, the classic $\langle e, t \rangle$ concept function with a domain of objects and a range of truth values becomes an $\langle s, \langle e, t \rangle \rangle$ function from possible worlds (‘s’ stands for a possible world) to functions from objects to truth values, analysable as we saw earlier in terms of sets of objects.

2.2.7 Identity as grammatical predication in Frege

It is highly relevant for this thesis that even identities are considered to be linguistic predications for Frege, the predicate being ‘*is* + object’ (meaning a logical object, not a grammatical one) where *is* is not a copula (i.e. a sign of predication) but stands for a concept, the concept at issue being—in the case of the sentence *The morning star is Venus*—‘that which is no other than Venus = The morning star’ (where *is* is a copula⁷). Frege gives exactly that concept function (Frege, 1892a, 44), noting, ‘admittedly only one object falls under this’. So for Frege, even identity statements were analysed in terms of concept functions. In fact, the concept for Frege in *The morning star is Venus* is really an $\langle e \langle e, t \rangle \rangle$ function, which would be represented in formal semantics as:

(31) $\lambda x. \lambda y. [y \text{ is no other than } x]$

(This makes it like a transitive verb such as *love*.) (31) is initially saturated by *Venus* before all of ‘is no other than Venus’ is saturated by *The morning star*.

There will be no disagreement from me in this thesis that identity sentences can be analysed as a kind of grammatical predication. The difference between what Frege said and what I argue later in chapters 4 and 5 is in how the *is* in *The morning star is Venus* gets characterised. I would like to hold on to the view that it is a copula in the same way as the *is* in *The morning star is bright*.

Hinzen (2016) is consistent with Frege’s view of identities as grammatical predication, although he goes further than Frege probably would in viewing this as a reason to consider them semantically property-ascription sentences of some kind:

⁷Jespersen (1924, 154) is, I think, consistent with Frege: ‘In the mathematical formula $A = B$ we should not take the sign $=$ as the copula and B as predicative, but insert the copula *is* before the predicative *equal to* B , and thus read it as meaning: A is comprised among the (possibly several) objects that are equal to B (whether ‘equal’ connotes only quantitative equality or perfect identity)’.

In [*Hesperus is Phosphorus*], ‘Hesperus’ is in a grammatically referential position. The sentential predicate is ‘is Phosphorus’... This predicate needs to pick out a property to apply to the subject. For this, it needs to exploit any descriptive information available. The only one available is: ‘being called Phosphorus’.
(Hinzen, 2016, 25)

Ultimately, the view that—as Hinzen (2016, 12) puts it—, ‘there is no logical relation of identity in grammar’ is one that I share in this thesis. And, in fact, I think that Hinzen hits on an important feature of identity statements (following Frege in the *Begriffsschrift*, as we’ll see in the next chapter) by highlighting the naming or ‘being called’ relation that they exhibit, in the quote above. However, I still think that it is worth considering whether the locution ‘applying the property of being called Phosphorus’ is helpful. It isn’t obvious that there is any substance to the notion that ‘being called *X*’ is a property (see section 7.3 for attempts from metaphysics to state a substantive notion of property).

Another move to make identities into more than just grammatical predications was Blom and Daalder (1977), who claimed that *The morning star is the evening star* predicates a property ‘being the evening star’ of the morning star, where *The evening star is the morning star* predicates ‘being the morning star’ of the evening star. But as Declerck (1988, 111) says:

[S]uch an analysis is quite counterintuitive. I discern no difference of meaning between *The Morning Star IS the Evening Star* and *The Evening Star IS the Morning Star*. Moreover, there is plenty of evidence that this type of sentence is different from predicational sentences.

This evidence includes: the postcopular nominal in an identity sentence is always a referring nominal, whereas the property nominal of a predicational sentence is by definition nonreferring; nuclear accent is on the copula in an identity statement (as the typography in the Declerck quote just given shows) where it is generally on either of the flankers in a predication sentence; the copula in an identity sentence can be paraphrased as *is the same as* or *is identical with* which is nonsensical in predicational sentences *My bicycle is blue* → *??My bicycle is identical with blue*; predicational sentences answer questions like *What is DP?* or *What is DP like?* e.g. *What is your bicycle like? It’s blue*. But this is not appropriate for identity sentences: *What is Hesperus like? #It’s Phosphorus*. So the prospects for treating identities as anything more than syntactic predications, as Blom and Daalder seem to want to do, don’t appear very promising, and I shan’t pursue that approach here.

2.2.8 Conclusions

So in this section, 2.2, we have followed through Frege's meaning of a predicate in terms of a concept, a particular kind of function with an unsaturated argument and a range set of truth values.

What matters here is that it is Frege's notion of a function as a root to the meaning of 'predicate' which has informed linguistics, especially formal semantics. Is there anything substantively meaningful about functions? It's not obvious that there is. Frege sometimes talks as though there are meaningful relations, e.g.

In the sentence 'Two is a prime' we find a relation designated, that of subsumption... This creates the impression that the relation of subsumption is a third element supervenient upon the object and the concept. This isn't the case ... the object engages immediately with the concept without need of special cement. Object and concept are fundamentally made for one another, and in subsumption we have their fundamental union. (Frege, PW, p.178, cited in Wiggins (1984, 327))

But here, it seems he has raised the spectre of a meaningful relation only to deny that it is really there. Where, exactly, Frege would argue subsumption comes from is not clear. It evidently isn't part of the logical representation of the sentence.

Functions appear to be part of the combinatorics of predicate logic, moreso than part of its semantics. As such, viewing predication—understood as functional application—as a meaningful term with a meaning equivalent to 'property ascription' seems like a category error. Property ascription may be a metaphysical notion rather than a logical one. Functions—mappings between two sets—aren't obviously meaningful in themselves. Of course, the set elements—understood as the denotations of constant symbols (or variable symbols under an assignment)—that they map from and to are though. This last point is essentially the idea (stated e.g. in Roussarie (2006)) that the meaningful elements of predicate logic are those given by an 'interpretation function'. An interpretation function of a formal language L specifies what each constant of L refers to, i.e. what entity in the domain of discourse D of L each constant refers to. So it is a function with the set of constants in L as its domain and the set of entities D as its range (Gamut, 1991b, 88). The main way of dealing with the interpretation of sentences where a quantifier binds a variable, like $\forall x\varphi$, is to use assignments, which are just functions which have the set of all variables of L as their domain and entities in D as their range (Gamut, 1991b, 96). In so far as domain and range sets of all functions are reducible to objects and truth-values ultimately, we have entities that are

meaningful (in so far as objects and truth-values are essentially metaphysical notions). But to consider a function itself as meaningful doesn't seem right.

In the other main part of this chapter, we shall switch topics and look at some of the ways that functions have entered into syntactic theory, and a syntactic definition of predication based on them. The main reason for doing this is to show that—similarly to logic—syntax doesn't obviously offer us an account of predication that could serve to distinguish some copular sentences from others. Indeed it isn't obvious that syntax has a problem-free account of predication.

2.3 Functions in linguistics, mainly syntax

It hasn't perhaps been recognised how fundamental functions are in linguistics generally, not just semantics. The point of this section is to show that they enter into syntax in several places. In particular I look at the way that certain kinds of function have entered into an attempt to define predication in syntactic terms. This, then, is another way that we might consider whether predication receives a meaningful 'property ascription'-like definition. Admittedly here the assumptions are somewhat more tendentious than in the case of logic determining meaning, though I think it is quite mainstream in syntax to believe that syntactic facts determine semantic ones, something which is often not explicitly stated, but is rather a common working assumption (though see the quotes from den Dikken (2006b) and others in the introduction to this chapter, section 2.1, for some relatively explicit statements). Beyond the way that this assumption may be unwarranted, there is a 'syntax-internal' difficulty to defining predication in syntactic terms that we shall encounter, namely that there is no problem-free such account. In the rest of this section, we shall build up to a definition of predication in syntactic terms, as external theta role assignment, 2.3.3, or simply external argument assignment—2.3.4, showing that it is not unproblematic. 2.3.5 then confirms this with Moro (1997)'s doubts about the possibility of eliminating predication as a primitive of linguistic theory.

The last two sections (2.3.6 and 2.3.7) do not directly build on sections 2.3.1 to 2.3.5. Section 2.3.6 mentions another function, type-shifting (which is, in fact, more a part of semantics than syntax) because I refer back to it later on in the thesis, in chapter 8. Section 2.3.7 is included in order to show how the central combinatorial operation of syntax—Merge—can also be viewed as a function.

2.3.1 Theta roles

So as we've seen, Frege's concepts are at the root of contemporary formal semantics. But there are also aspects of syntax that deal in unsaturated functions. As Roy (2013, 5) says:

A predicate is a lexical item denoting an open function and whose open positions must be completed by an argument. Thus, verbs, for instance, have a theta-grid that specifies lexically their thematic properties, and that must be saturated by closure by an argument.

Theta roles are worth exploring here. The reason for this is twofold: Firstly, theta roles are kind of function. Secondly, we can try to see whether predication might be reducible to, or substantially explained by, the fact of a syntactic object being a theta role assigner. As den Dikken (2006a, 10) puts it:

In present-day linguistic theory, what we have is a notion of n -ary predicates (ultimately based on Frege), expressed in terms of theta-grids, in conjunction with the Fregean notion of "saturation,"... A lexical item with a theta-grid is called a predicate, and so, usually, is the combination of the predicate head and its internal argument(s)—that is, the syntactic constituent that is predicated of the subject.

Theta roles are a formalisation of the intuition that part of any given verb's meaning relates to the participants in the activity it describes. Haegeman and Guéron (1999) take the example of *abandon*. It's hard to think of this activity without thinking of the person who abandons and the person or thing who is abandoned: 'the verb requires a number of participants to engage in a certain state of affairs' (25):

(32) Thelma abandoned the project.

In terms of their theta roles, *Thelma* is the AGENT, *the project* is the PATIENT. (Theta roles are typically put in small caps.) Other theta roles used in the literature are (list from Haegeman and Guéron, 26):

- (33)
- a. AGENT/ACTOR: the one who intentionally initiates the action.
 - b. PATIENT: the person or thing undergoing the action.
 - c. THEME: the person or thing moved by the action.
 - d. EXPERIENCER: the entity that experiences some (psychological) state.
 - e. BENEFACTIVE/BENEFICIARY: the entity that benefits from the action.
 - f. GOAL: the entity towards which the activity is directed.

- g. SOURCE: the entity from which something is moved as a result of the activity.
- h. LOCATION: the place in which the action or state is situated.

For example:

- (34) a. Thelma handed Louise the text
 AGENT verb BENEFICIARY THEME
- b. Thelma liked the text
 EXPERIENCER verb THEME

An important condition on the well-formedness of a sentence is known as the ‘theta-criterion’. This basically makes the point that there is a one-to-one correlation between theta roles and the arguments that realise them in a sentence, and furthermore that an argument must realise a theta role. Therefore, (35) is bad because there are three arguments in the sentence for a verb that only assigns two theta roles:

- (35) *Thelma abandoned the book the project.

In (36) the verb *handed* assigns three theta roles but there are only two arguments in the sentence, and so the sentence is ill-formed because there should be three arguments. Alternatively, the badness of this sentence can be understood as a result of one argument taking on two theta roles, something else which is prohibited by the theta-criterion:

- (36) *Thelma handed Louise.

(37) **Theta-criterion**

- a. Each argument must be associated with one and only one theta role.
- b. Each theta role must be associated with one and only one argument.

The theta-criterion can guide us in assigning structure to sentences. For example, sentences like

- (38) Thelma believes Louise to have abandoned her husband.

which semantically approximate

- (39) Thelma believes the story.

more so than

- (40) Thelma believes Louise.

can be analysed as structures where *Louise* is assigned a theta role by *abandoned*, not by

believes. So *believes* assigns two theta roles (to *Thelma* and the whole constituent *Louise to have abandoned her husband*) and *abandoned* assigns two theta roles (to *Louise* and *her husband*). The fact that one argument can only have one theta role guides us to that conclusion. Otherwise, we might prefer an analysis where *Louise* receives theta roles from both *believes* and *abandons*. But the theta-criterion rules this out.

So much for the basic idea of theta roles. As I said, they can essentially be understood as functions, as unsaturated variable slots in a verb that must be saturated by an argument. Shortly, in 2.3.3, I'll look at an influential account that has attempted to base a definition of predication upon theta roles in order to see whether syntax can provide us with a notion of predication that can help explain the copular taxonomy. In order to understand that theory, we need to briefly look at the basic idea of the internal–external argument distinction in syntax.

2.3.2 Internal and external arguments

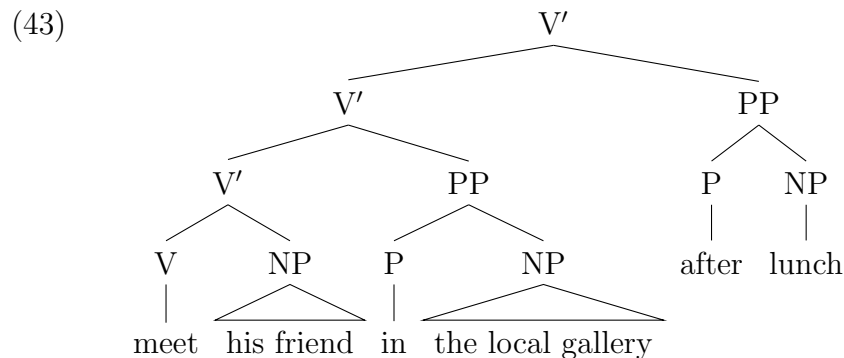
The internal–external argument distinction is the second piece of the puzzle we need to understand the definition of predication given in the next section, 2.3.3. Internal and external arguments can also be understood as functions. This originated with the analysis of the verb phrase in terms of layers between the root verb beyond the maximal VP layer, i.e. with an analysis that includes intermediate projections (I'll say a bit more about this in the footnote 9). Take the following pattern of ellipsis (from Haegeman and Guéron (1999)):

- (41) a. The young linguist will meet his friend in the local gallery after lunch.
 b. He will do so indeed.
 c. He will do so before lunch.
 d. He will do so at the university before class.

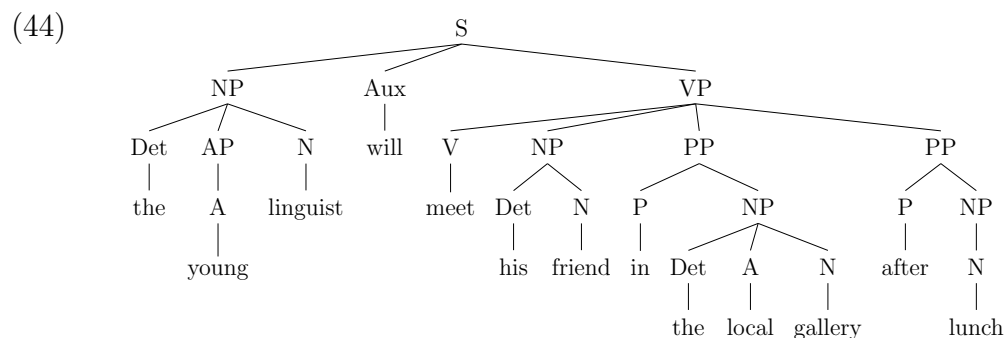
Following the well-established view that only complete syntactic constituents can be ellided (rather than, say, half a constituent), this indicates the following strings of (41a) are syntactic constituents that should appear as such in any analysis we give of (41a):

- (42) a. meet his friend in the local gallery after lunch
 b. meet his friend in the local gallery
 c. meet his friend

The following tree captures these constituency facts:



This is not captured by a traditional analysis (which is also not binary branching and so not in line with mainstream phrase structure analysis since Kayne (1984))⁸:



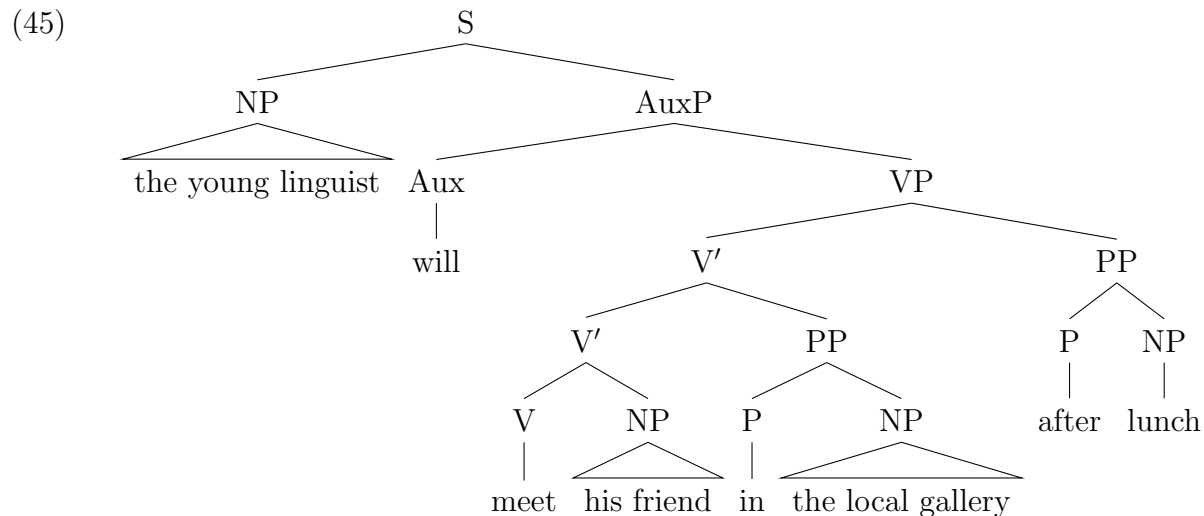
Note that the strings *meet his friend* and *meet his friend in the local gallery* are not constituents in the tree in (44), so it is easy to see that this second tree does not contain constituents corresponding to those in (42).⁹

An argument dominated by V' is an **internal** argument. So that includes the NPs *his friend* and *the local gallery* in (42). The argument (there is only ever one) that is realized outside the VP is called the **external** argument. If we complete the structure of (43), we can see that *the young linguist* is the external argument:¹⁰

⁸'Aux' stands for 'auxiliary verb'.

⁹ So, going back to the remark I made about layers between the root verb and the maximal VP layer: all this means is that traditionally syntax analysed word classes in terms of root categories and then phrases which contained these root categories. We can see that in (44); there is just a root verb (V) and a verb phrase (VP) that contains the V *meet* as well as *his friend in the local gallery after lunch*. The basic idea—which is absolutely fundamental to syntax—is that it is the verb which 'projects' in some way, such that we can say that all of *meet his friend in the local gallery after lunch* is verbal rather than nominal or prepositional. If it was nominal, for example, we should expect to be able to preface it with an article, but we can't (**the meet his friend in the local gallery after lunch*); we could if the phrase was *meeting of his friend in the local gallery after lunch*, however, so in this case we would call *meeting* a noun (N) and the whole phrase an NP. In any case, it was soon felt, on the basis of such evidence as the constituency tests given in (42) that there was a need for an intermediate layer between the root X and the phrase XP, which is notated X', and which we see in (43). Chomsky (1970) is a classic source for this 'X'-theory', subsequently elaborated by Jackendoff (1977).

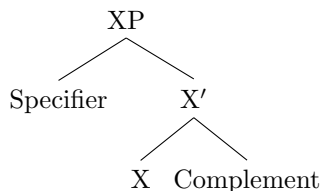
¹⁰The adjunct PP *after lunch* should—like adjuncts generally in Government and Binding Theory—be adjoined to a zero- or bar-level phrase (i.e. the head X⁰ or X') and project a same-level phrase. (45),



A final point to make about the internal/external argument distinction is that whilst [Spec, IP] is often considered to be the canonical subject position in configurational terms, it is now widely assumed that this is a derived position, and that the NP which ends up in that position is in fact base-generated internally to the VP.¹¹ This VP-internal subject hypothesis is motivated by the behaviour of floating quantifiers (see Haegeman and Guéron (1999, 228–31)), as we see in the following examples, where the quantifiers *both* and *all* can apparently

therefore, is not quite correct; there should at least be a V' dominating the PP *after lunch*.

¹¹ [Spec,IP] means 'the specifier of IP'. So what is a specifier? And what is IP? X'-theory essentially claimed that all syntactic structure is built out of phrases with a uniform structure (this is the 'core' phrasal structure, before the addition of adjuncts):



This was just an empirical finding about phrase structure (see footnote 9). 'IP' (for 'inflection phrase') was in many ways the successor of 'S' (for 'sentence') in the Government and Binding theory of UG (Chomsky, 1986), the highest node in a sentence. The problem with S was that it made syntax 'exocentric' (as pointed out by e.g. Emonds (1976)), meaning that it contained categories at a higher level that were not of the same type as the categories they hierarchically dominate. (44), and indeed (45), are exocentric as there is a category 'S' which does not appear lower down in the tree; in other words, the sentence as a whole is not a projection of either a noun (NP), an auxiliary verb (Aux), or a verb (VP), S's immediate daughters in (44) and (45). Chomsky (1986)'s introduction of the 'inflection' (I) category (to somewhat replace the 'Aux' of (44) and (45), and generally to make sentences essentially verbal projections) and his making the highest node the maximal projection of I—IP—made syntax endocentric, a desirable result for theoretical and empirical reasons. In fact, the 'canonical subject position' is nowadays generally referred to as Spec,TP ('TP' for 'tense phrase') (Biberauer, 2004, 15), following the demise of IP initially into distinct inflectional phrase heads such as TP and subject and object agreement phrases (AgrSP, AgrOP) to capture agreement inflections; the latter subsequently were superseded by phi-features as properties of lexical items themselves in the Minimalist Program (Chomsky, 1995), though TP remains a commonly used phrasal category. Later in chapter 8, I'll refer to the canonical subject position as Spec,TP reflecting the more modern syntactic apparatus I'll be referring to there.

‘float’ to the right:

- (46) a. [All the girls] have bought the book.
 b. [The girls] have all bought the book.
- (47) a. [Both my daughters] are studying linguistics.
 b. [My daughters] are both studying linguistics.

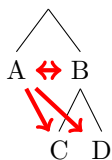
We would like an analysis of these sentences that reflects the fact that they are semantically parallel. An analysis in terms of rightward movement of the quantifier would infringe the well-established requirement that moved constituents always c-command their traces.¹² (A ‘trace’ of movement is the putative position from which a constituent is moved by a syntactic operation. In structural descriptions, traces are often subscripted with letters in order to help keep track of what they are traces of, where there are multiple traces.) In the cases in (46) and (47) the trace would c-command the moved constituent. So we are forced to adopt another analysis. Under this analysis, the quantified NP *all the girls* originates at the point where we see the *all* in (46b):

¹² C-command is the most fundamental relation in syntax (‘[C]ertain—perhaps all—syntactic relations are restricted by c-command’ (Brody, 2002, 29)), ultimately originating in work by Langacker (1969). It says:

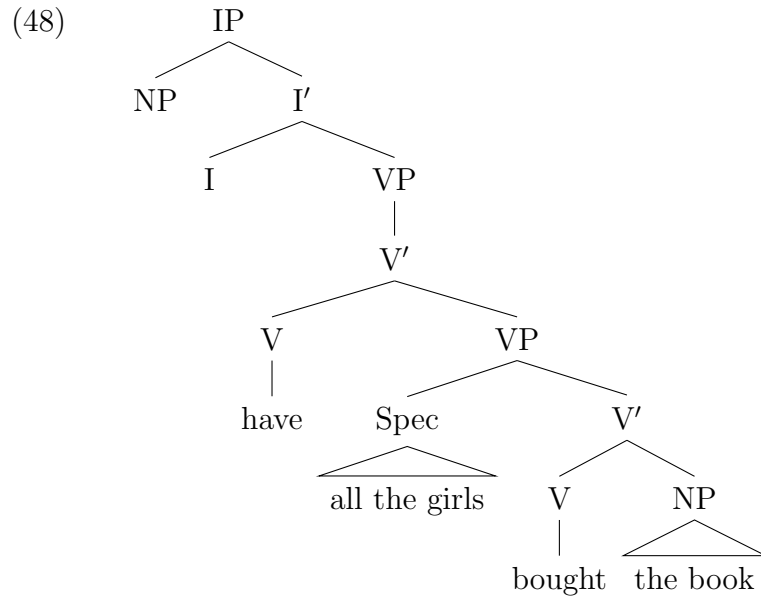
A c-commands B iff:

- a.) The first branching node dominating A dominates B, and
 b.) A does not dominate B, and
 c.) A does not equal B

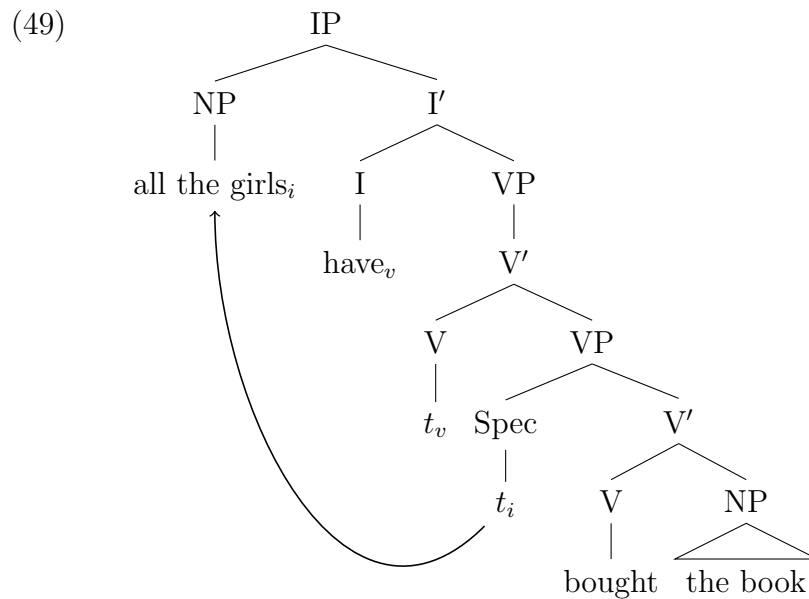
This definition (essentially Reinhart (1979)’s) seems complex, but it can really be understood diagrammatically very simply:



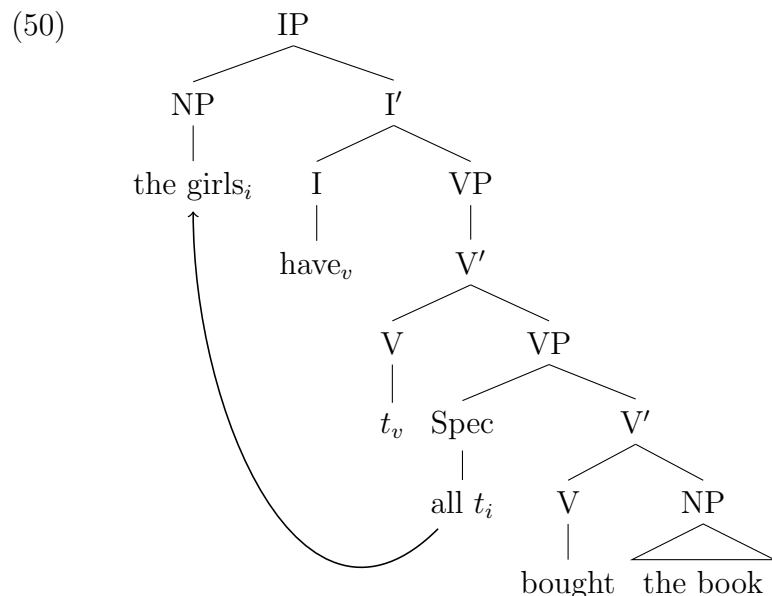
Here, A c-commands B, C, and D i.e. its sister and everything its sister dominates. C-command can generally be worked out like this: take a node, go up one node, then everything dominated by that node will be in the c-command domain of the original node. Note that B also c-commands A, according to the definition given above, and so A and B—like sisters generally—are in a relation of ‘mutual c-command’.



Then in order to derive the two sentences in (46), either the whole of *all the girls* moves up the tree to the canonical subject position [Spec, IP], or just *the girls* moves, leaving *all* in situ in [Spec,VP]. In other words, we either have the following derivation for (46a):

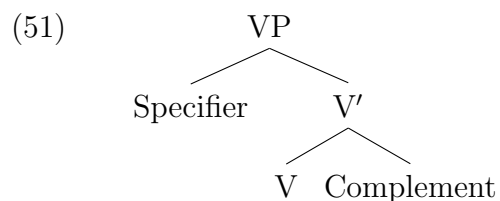


or we have the following derivation for (46b) (I shan't go into the details of what factors affect whether we end up with (46a) or (46b)):



So, to conclude this section, syntax appears to make a distinction between internal and external arguments available, whereby an external argument appears outside VP and an internal argument appears inside VP, dominated by V'. This can either be a result of the base-generated positions of the arguments, or the external argument can move to the external position, such that it is a derived position.

How can the internal/external argument be understood in terms of functions? Rather like a theta-grid, we can think of a verbal structure as something with placeholder slots corresponding to the number of arguments it takes, its 'valency'. So, for a transitive verb with one direct object there are two argument slots and the schematic structure would be:



Given two arguments, one is mapped to the complement and is thus the internal argument; the other is mapped to the specifier (from which it moves to a position outside VP, such as [Spec,IP]—the canonical subject position) and is thus the external argument. This is a considerable over-simplification with respect to modern syntactic theory, where the external argument has been argued to be introduced outside the verb phrase altogether e.g. by a Voice head (Kratzer, 1996), but it will serve for our purposes to merely understand that there is a difference between a verbal argument which is c-commanded by the V head itself,

and one that isn't.

2.3.3 Subjects as external theta roles

Stowell (1981; 1983) aligns the predicate with the X' of X-bar theory (see footnote 9) and the subject with the specifier of XP; tensed clauses are considered projections of Inflection (I) such that Spec,IP is the subject of a sentence and small clauses have a parallel structure except without an inflectional element. In the case of sentences where the small clause was headed by a verb as in (52a), Stowell therefore postulated a subject position inside VP, which the subject raises out of in other cases, (52b,c):

- (52) a. Jane saw [_{VP} the students riot]
 b. [_{IP} Jane_i [_{VP} t_i saw Max]]
 c. [_{IP} Max_i [_{VP} t_i danced]]

As Svenonius (1994, 16) says, this paved the way for the VP-internal subject hypothesis (along with the sort of evidence we saw from floating quantifiers in section 2.3.2) and made for a conceptually simpler system of theta role assignment, where all theta roles could be assigned within the maximal projection of a theta assigner (the verb in (52a)). This system relies on theta role assignment for predication to arise. On this view, predication is not fundamentally different from the relation between a verb and its direct object. Direct objects must also receive a theta role and have a specific structural configuration (i.e. they must be the complement of XP). However, there may be reason to consider that a special theory is required to treat subjects as distinct from objects (rather than treating predication, as Stowell—and the Government and Binding theory of Universal Grammar did generally—as essentially just argument-taking, except for a difference in the surface structure in the syntactic tree). Svenonius (1994, 25–6) points to the existence of ‘subject–object asymmetries’: objects can sometimes be ellided where the context allows a generic/indefinite sense for the object e.g. *I wrote, I spoke, I painted, I lifted*. Subjects, however, cannot generally be so omitted. Svenonius gives the example of a ‘causative–inchoative alternation’ *The captain sank the boat to collect the insurance* vs. **The boat sank to collect the insurance* (the alternation is ok when it is just *The captain sank the boat* vs. *The boat sank*), saying ‘a missing causer is not contextually supplied but presumed nonexistent’ (Svenonius, 1994, 26).

A variation on the GB/Stowell theme is a position advocated in a series of papers by Williams (1980; 1983a; 1987), about which Rothstein says he was ‘making implicit use of Frege’s classification of linguistic objects into saturated and unsaturated’ (Rothstein, 2006, 74). Technically, for Williams, the predicate is XP (not X' , as it is for Stowell) and it

assigns an external theta role via coindexation under mutual c-command, which means that the argument assigned the external theta role (i.e. the subject) is, perforce, an external argument of the predicate in the sense of the last section, 2.3.2. Williams himself allowed greater-than-binary branching, though as we mentioned in the last section this is not in line with modern syntactic theory. The internal–external theta role distinction is essentially taken as basic; Williams has a system whereby verbs (and all other types of predicate/theta role assigner) have an argument structure, where one argument will be the ‘head’ (Svenonius, 1994, 18). So predication, for Williams, is the assignment of an external theta role (in his technical sense), which will necessarily be an external argument in the sense of the last section, 2.3.2. For simplicity, then, we can abstract from Williams’s technical notion of an external theta role (i.e. the head argument in a theta-assigner’s argument structure) and simply think about a subject as an ‘external theta role’ in the sense of it being the theta role that is assigned to the external argument in the sense of the last section, i.e.

- (53) External theta role = the theta role that is assigned to the external argument (the argument that either originates or moves to a position outside the maximal verbal projection)

Could we then say that a subject is defined as an ‘external theta role’ of a predicate, and (co-dependently) that a predicate is something that assigns an external theta role to a subject?

Now, it so happens that being the external theta role is commonly found as a property of subjects. Rothstein (2001) gives the example:

- (54) Mary ate the carrots raw.

In (54), *raw* assigns an external theta role to *carrots* and *eat* assigns an external theta role to *Mary*.

However, there are some immediate empirical problems with aligning external theta roles with subjects which Rothstein (2001) identifies. Firstly, there are cases where the subject is not assigned an external theta role. The most obvious instance of this is a passive, like:

- (55) $Mary_i$ was [haunted t_i by a ghost].

Here, the subject *Mary* is assigned an *internal* theta role by *haunted* in the position of its trace.

Another case is where the subject receives an external theta role, but it isn’t that of the matrix verb. This is the case in a raising construction, like:

- (56) $Mary_i$ seems [t_i to like the ghost].

Here, *Mary* gets an external theta role, but it isn't the external theta role of *seems* but rather of *like* in the embedded clause. In so far as there is an intuition to consider the 'main predicate' of (56) *seems to like the ghost*, then this would not tally with a definition of the predicate as the external theta role assigner.

Interestingly, Rothstein (2001) also gives examples of equatives like:

(57) John is Mr Smith.

Here, there is a 'non-predicative complement' and '*be* itself doesn't assign a theta role'. So, again, these are counter-examples to the view that predication is external theta role assignment. About such examples, Rothstein says (62–3), '...the copular does not take a predicate as a complement, so that the VP does not inherit from its complement an external theta role to assign. They thus exemplify a kind of non-thematic predication, since the VP predicate is headed by a copula which has an argument complement and no external theta role to assign'. Its acceptability, for Rothstein, is because 'V projects a predicate node', and this is in line with Rothstein's general view of predication as a purely syntactic relationship i.e. not involving theta roles; Rothstein considers theta roles a lexical semantic affair.

Here is another example (due to Moro (1997)) where a subject appears to receive no theta role:

(58) I consider that John's fear of elevators.

Here, *fear* has assigned all its theta roles to *John* and *elevators*, in such a way that—whilst *that* is still a subject of predication, it is not obviously assigned any theta role.

Finally, there is the case of pleonastic subjects, what Rothstein (2001, 65) calls 'the most obvious example of subjects which are not thematically related to their predicates'. Examples of these are:

- (59) a. I believe it to have turned out that John won the race.
 b. Il a été mangé trois pommes. French
 it has been eaten three apples
 'Three apples have been eaten.'
 c. ze meatzben Se dani meaxer. Hebrew
 it irritates that dani is-late
 'It is irritating that Dani is late.'

Here we have both no theta role being assigned to the pleonastic subject *it/il/ze*, in fact not even 'any semantic predication relation' (Rothstein, 2001). Rather, the presence of *it* in this sentence appears to be a purely syntactic requirement, not reflective of any semantic requirement at all. Williams attempts to circumvent these problems (in the latter case of

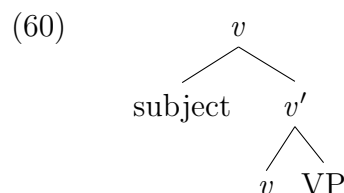
the pleonastic subjects, he simply denies that there is any predication relation involved), but Rothstein (2001, ch.4) argues convincingly against his technical solutions.

The upshot of this, then, is that it seems that there are good empirical reasons to doubt that predication can be defined in terms of a predicate being an external theta role assigner, as Williams (1983a) tried to argue. Ultimately, therefore, it seems that a purely syntactic definition of predication is hard to maintain.

2.3.4 Might subjects simply be the same as external arguments?

Whilst Williams has a notion of external argument, it is tied to theta-theory. It is the constituent (DP) that is assigned a theta role by a predicate and realised as its external argument.¹³ But as we saw in the last section, 2.3.3, this had empirical problems. However, there is of course the simpler notion of external argument of a verb that doesn't relate directly to theta-theory that we encountered earlier (section 2.3.2). Perhaps this notion is all that we need in order to define subject and predicate syntactically; let's explore this possibility.

External arguments of verbs have been argued to be introduced not by the verb itself (and are therefore no part of its thematic grid) but by a functional head which hierarchically embeds the lexical verb phrase (VP), little *v*, what Chomsky (1995, ch.4) calls a 'light verb' (Chomsky follows up work by Hale and Keyser (1993)). [Spec, *v*P], then, is the position which hosts the external argument, internal arguments appearing as the specifier or complement of VP.¹⁴

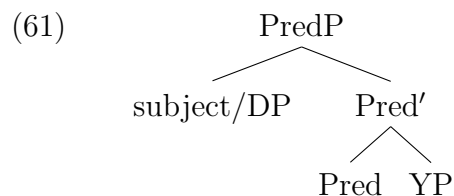


This gives us a syntactic definition of the subject of a verb. Let's see how this could extend to the subject of a nonverbal predication.

¹³'DP' stands for 'determiner phrase'. Many syntacticians believe that all noun phrases with articles (definite, indefinite, demonstrative, possessive etc.) are DPs, that is, projections of a determiner head which takes a noun phrase complement, rather than projections of the noun phrase itself. This view is known in Universal Grammar as the 'DP-hypothesis' (notably associated with Abney (1987)) and is theoretically very elegant in that it established a structural parallelism between nominals (headed by a functional category (D) with a lexical NP complement) and sentences (headed by a functional category (I) with a lexical VP complement). See section 4.3 for the distinction between functional and lexical categories in linguistics. Other syntacticians, like Roy (2013) believe that only *definite* noun phrases are DPs. It isn't always obvious which sense is intended.

¹⁴As mentioned a little earlier, Kratzer (1996) has suggested that this *v*P category is essentially a projection of voice (so, the phrase 'Voice Phrase').

For Roy (2013) (following Bowers (2001), Adger and Ramchand (2003), and the spirit of den Dikken (2006a)) nonverbal subjects can be defined structurally as the specifier of a specific ‘predication phrase’ (for den Dikken (2006a) it is a ‘relator phrase’ where several different constituents, e.g. copulas, verbs, and prepositions, may count as relators). This, then, means that there is a parallelism with respect to verbal and nonverbal subjects.



where YP can be AP (‘adjective phrase’), NP, DP, and PP (‘preposition phrase’).

These structures can be viewed as functions, with the constituent in specifier position acting as the saturator of the function. In the case of the nonverbal predication in (61), Bowers (1993) proposes that the complement of such a structure (the YP) is a property, a saturated constituent that is rendered unsaturated/predicative by the head Pred.

Where copular sentences are concerned, there is a notable difference between Roy and den Dikken with respect to the role of the copula itself. For den Dikken, the copula is one of the elements that mediates predication, i.e. that would be the Pred head of (61). For Roy, this is never the case; the copula is a raising verb that selects PredP as its complement (following Stowell (1978)). And so for Roy, Pred is always presumably an abstract head without a phonological reflex.

Such a purely syntactic definition of subject and predicate is also in the spirit of Rothstein (1983; 2001), and is quite a standard view of the syntactic representation of predication. It is, however, not terribly explanatory. Identity sentences like

(62) The evening star is Hesperus.

are almost universally considered *not* to be predications (den Dikken (2006a) is an exception; he explicitly analyses equative copulars in terms of his relator phrases). In syntactic terms, then, they would not contain a PredP. But why not? And if there is no PredP in (62), then it is hardly explanatory to say that we can explain why

(63) The evening star is bright.

is a predication on the basis of it containing a PredP which we have stipulated to be there. In many ways, PredP is as useful as the notion of ‘property ascription’; anything we want to be PredP can be one. This is because it is just an instance of the X-bar schema, and every sentence is analysable in terms of X-bar schemata. Whatever a subject is, it can certainly

be analysed as a specifier of an X' constituent. In English, this is plausible because subjects always precede their predicates. In languages where subjects are not pronounced (say, in Italian, where there may only be agreement-marking on the verb), or in languages where subjects follow the verb (say, verb–subject–object languages like classical Arabic and Irish Gaelic), syntax has to resort to two mechanisms: 1.) the mechanism of invisible constituents like *pro*, which occupy the external argument/[Spec,*v*P] position. 2.) the mechanism of invisible movements of the subject either from or into the external argument/[Spec,*v*P/PredP] position. Evidently, such mechanisms risk the theory becoming unfalsifiable, and so such proposals have to be judged carefully with respect to what other empirical work they do. This scepticism about PredP being a part of the functional categories of Universal Grammar is in line with recent work by Marelj and Matushansky (2015) and Matushansky (2015; to appear) who draw on substantial empirical evidence to reach this conclusion.

2.3.5 Eliminating predication?

Moro (1997) is in many ways an attempt to discover whether the term ‘predication’ can be eliminated altogether from syntactic theory. According to Moro, this was considered to be a typically semantic notion, which, *prima facie*, syntax shouldn’t deal with. Moro (1997, 77) talks about

... a basic assumption of generative grammar (in fact, an assumption which is common to all distributionalist theories), namely that linguistics should avoid ‘semantic’ notions such as predication.

But Moro nicely spots the way that doing this came at a cost of missing how noun phrases could function as predicates. He notices that this is implicitly ruled out in a passage from Chomsky (1993, 517):

For concreteness, take categories to be as in (14), for nouns, verbs, adjectives, and pre-/post-positions, respectively

(14)

(i) N = [+ N, – V]

(ii) V = [– N, + V]

(iii) A = [+ N, + V]

(iv) P = [– N, – V]

The feature [+N] is the traditional substantive; the feature [+V], predicate.

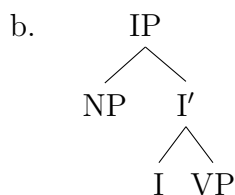
The rest of Moro (1997) is an attempt to show that it is not possible to entirely eliminate the notion of predication and explain the facts about inverse copular sentences (sentences

where the base-generated flanker moves to the front of the sentence, as is commonly agreed to happen in specificational copular sentences) purely in terms of such structural features.

Moro notices that a major assumption of Universal Grammar is that the ‘skeleton’ of the clause, as he puts it, is given by the inflectional morphemes, which are essentially verbal agreement and tense. A second assumption is that what Moro calls ‘grammatical functions’ such as ‘object’, ‘subject’, and ‘predicate’ should be ‘entirely derived from syntactic configurations’. What he then proceeds to show is that these two assumptions encounter problems, problems which cannot be solved unless we do in fact allow that ‘subject’ and ‘predicate’, at least, are still required as explanatory primitives of syntactic theory.

The first assumption could correspond with X' -theory. That is to say, X' -theory is a possible implementation of the idea that the clause has a fixed skeleton, based on the nature of inflectional morphemes. This X' -structure of the clause is: the head is I (containing agreement and tense inflectional morphemes), the complement of I is VP, and the specifier of I is NP. In other words,

(64) a. $[_{IP} NP [_{I'} I VP]]$, or



Moro calls this the ‘predicative nucleus’ of a sentence. Based on this, we can define grammatical functions in purely configurational terms:

The subject is the NP which is immediately contained by the clausal node, IP;
 the object is the NP which is immediately contained by the VP. Eventually, the
 predicate is the VP itself.

When this framework is confronted with copular sentences, various problems arise, notably the problem whereby the following two sentences are considered variants of each other:

- (65) a. There is a man in the room.
 b. A man is in the room.

Moro asks ‘Why is this so?’, meaning: Why is it the case that if *a man* does not end up in pre-copular position at surface level do we need to insert an expletive *there*? There is no ‘valency’ requirement (or theta role satisfaction requirement) here, as the fact that (65b) is acceptable attests. It is not like the way that **John likes* or **Likes movies* is bad because

an argument that the verb *likes* requires is missing, violating the theta-criterion (see section 2.3.1).

The solution (or, perhaps, descriptive generalisation) to this problem (the problem of why we need to insert an expletive *there* in (65b)) that Chomsky (1986, 116) proposes is known as the ‘extended projection principle’, and as Moro again notices, it crucially makes reference to the notion of subject of a predication:

The projection principle requires that complements of heads must be represented at each syntactic level (D[EEP]-structure, S[URFACE]-structure, L[OGICAL] F[ORM]), so that, in particular, objects must be represented, but it says nothing about subjects. Thus, it distinguishes between what Edwin Williams calls “internal” and “external” arguments, specifically, object and subject. The projection principle requires that the former be syntactically realized, but not the latter, although they are required as subjects of predication (either arguments or expletives). The two principles—the projection principle and the requirement that clauses have subjects—constitute what is called the extended projection principle (EPP) in Chomsky (1981). (Chomsky, 1986, 116)

As Moro says, ‘... [t]he theory cannot rely solely on configurations’; the EPP clearly resorts to a primitive notion of ‘subject of predication’.

Interestingly, the two clauses of the EPP are explicitly related in Fregean terms by Chomsky. Referring to Rothstein (1983), he says:

We may think of a lexical head as a “lexical function” that is “unsaturated” (in roughly the Fregean sense) if it is not provided with appropriate arguments fulfilling the theta roles it assigns, and we may correspondingly regard a maximal projection (apart from those that are quasireferential: NP and clause) as a “syntactic function” that is unsaturated if not provided with a subject of which it is predicated. Then, the EPP is a particular way of expressing the general principle that all functions must be saturated. (Chomsky, 1986, 116)

Moro concludes, therefore, that where traditional notions like ‘active sentence’, ‘relative clause’, ‘intransitive verb’ have all been given syntactic analyses that have broken them down into ‘more abstract entities’, such that their continued use is a ‘purely taxonomic’ affair, this is not the case for ‘subject’ and ‘predicate’: ‘[T]hese two notions still appear to play a substantial part in syntactic explanation’ (Moro, 1997, 261).

So we have seen in this section 2.3 that it is not obvious how we can give a syntactic definition of predicate. The view that a predicate is an element assigned a theta role which is also the

external argument of the theta role assigner faces empirical problems, as we saw in 2.3.3. The view that predicates are simply a *sui generis* function head ‘Predicate Phrase’ we found to be too stipulative, and so not explanatory. Adding Moro’s conclusion, just noted, to this picture appears to lead to the following situation: ‘Subject–predicate’ has not been eliminated from the explanatory apparatus of syntax. And yet, a definition of predication in syntactic terms which is free from empirical problems has not obviously yet been achieved.¹⁵ So ‘subject’ and ‘predicate’ must apparently remain in syntax as unexplained primitives.

2.3.6 Type-shifting

This section and the next are examples of two other areas where functions have entered into linguistic analysis (more so in semantics than syntax, in fact, in this section). I refer back to type-shifting later in the thesis, chapter 8, so I’ll mention it briefly here. Type-shifting rules have been a part of the formal semantic toolkit since at least Partee (1983). Partee (1986b) uses them in the context of noun-phrase interpretation. Type-shifting rules are not ‘a single uniform and universal set’ (208); there are even type-shifting rules that are language particular. Type-shifting has notably been introduced in the nominal domain, and is really just a method for formal semantics to capture the ambiguity of noun phrases, the fact that e.g. a name like *Einstein* can appear to denote an object in some sentences, but something more like a quantificational phrase in others, such as when we say *He’s a real Einstein*, essentially turning it into a common noun. Chierchia (1985) also used type-shifting to account for the way that properties can be predicated of other properties:

- (66) a. John is wise.
 b. Bill is silly.
 c. Being wise is silly.

In (66a,b) *wise* and *silly* are normal property-denoting (for Chierchia) $\langle e, t \rangle$ functions. But *being wise* in (66c)—in so far as it expresses a property—is also arguably an $\langle e, t \rangle$ function. In that case, *silly* in (66c) couldn’t also be an $\langle e, t \rangle$ function—as it is in (66a,b)—as semantic composition between two ‘same level’ functions is impossible; one function must always be a higher level function. We can either claim that adjectives are systematically lexically ambiguous, or (Chierchia’s preferred solution) we can postulate an operation, or set of operations, that can apply to functions generally, shifting them to higher or lower level functions (in this case from $\langle e, t \rangle$ to $\langle \langle e, t \rangle t \rangle$).

¹⁵Rothstein’s work e.g. (1983; 2001), the details of which would take us too far into syntactic complexity, is a substantial attempt to define predication syntactically.

These type-shifting operations are functions. This is revealed by Partee (1986b) using the classifications of functions for them. For example, one type-shifting operation called ‘lift’ shifts $\langle e \rangle$ -type expressions to $\langle \langle e, t \rangle t \rangle$ -type expressions. Partee says that this is ‘total and injective’, terms that are used to describe functions. A total function maps every element of its domain onto an element of its range. An injective function maintains distinctness, such that no element of the domain is mapped onto more than one element of the range. In other words, in a total and injective function, there is a one-to-one mapping between the domain and the range, with the domain being exhausted (but not necessarily the range).

One type-shifting operation which is relevant for this thesis is ‘ident’:

$$(67) \quad \text{‘ident’}: j \rightarrow \lambda x.[x = j] \text{ (total; injective)}$$

This function is often given as the meaning of the syntactic predicate of an equative copular sentence, and according to Partee (1986b, 210), it is an operation ‘mapping any element onto its singleton set’ (remembering that formal semantics translates functions into sets).

Chierchia (1984) defines two type-shifting operations ‘nom’ and its inverse ‘pred’. ‘Nom’ takes properties, $\langle e, t \rangle$ -type expressions, and maps them onto $\langle e \rangle$ -type expressions. ‘Pred’ does the reverse. ‘Nom’ is used, among other things, to show how the common noun *dog* can receive a mass-like reading in the bare plural *dogs*, or the way the adjective *blue* in *The car is blue* can become something like a proper name, as in *My favourite colour is blue*.

Type-shifting, then, seems to be essentially concerned with the issue of the referential ambiguity of lexical items, notably in the nominal domain. The same phenomenon has been treated in other theoretical frameworks. Sheehan and Hinzen (2011) and Martin and Hinzen (2014), for example, view referentiality as essentially following from syntactic structure-building operations, in particular they argue referential specificity increases in line with the ‘phase structure’ of recent Universal Grammar research (e.g. Chomsky (2001)). It is perhaps intuitively obvious that this is the case. After all, to take one of the examples above, *dogs* is not the same word as *dog* and we can expect the addition of the plural morpheme—understood as part of the syntactic component—to make an interpretive difference.

2.3.7 Merge as a function

For McGilvray, (Chomsky & McGilvray, 2012), science generally resorts to functions to state its explanations; indeed, it is part of being a properly scientific endeavour that it is reducible to functions. McGilvray thus views the central compositional principle of UG—Merge (or, to be technically more precise, ‘external Merge’)—as a function that takes two lexical items from its domain of values and maps them to a set consisting of them both: ‘X merged with

Y yields $\{X,Y\}$ ’ i.e. the operation Merge takes X and Y from its domain and maps them to $\{X,Y\}$ in its range. A second operation ‘internal Merge’ ‘comes for free’ according to Chomsky, given external Merge, and is the operation whereby a copy of an element inside a complex lexical item— $\{X,Y,Z\}$, say—is merged with that item—forming $\{Y,\{X,Y,Z\}\}$. This operation is important as it captures the way that natural language universally appears to allow for ‘displacement’ of lexical items from the positions where they were originally (externally) merged in a derivation. In fact, from this we can see that the displacement is somewhat illusory. What actually happens is that a copy is merged, which copy is then the only one that actually gets pronounced.

To give a concrete example of this (abstracting from lots of details of phrasal heads like tense, aspect, mood etc.), in order to get the sentence *What did John eat?*, we start by externally merging *eat* and *what*. We then externally merge *John* to *eat what*. We then ‘internally merge’ a copy of *what* to *John eat what* to make *What John eat what*, where we only pronounce the copy: *What John eat ~~what~~*. It is then just a detail of English that such a structure requires the insertion of an auxiliary verb *did* to carry the tense features: *What did John eat?* ‘Internal Merge’ used to be called ‘Move’, which reflected the way that models of UG prior to the Minimalist Program considered that lexical items literally moved in some sense, that *what* literally moved to the front of the sentence. Now, as we’ve just seen, the current view is that a copy is merged. It isn’t even obvious that ‘internal Merge’ felicitously describes this operation. Why should the ‘copy’ *what* be any different to the original *what* merged with *ate*? Syntactic theories like Brody (1995) have thus considered that everything is ‘externally merged’, that the ‘copy’ *what* is taken from the lexicon in the same way as the ‘original’ *what*.

An ‘extensional definition’ of Merge is not available, given that Merge is a recursive function, and so there are an infinite number of potential ranges. This makes it like the successor function of mathematics that produces the natural numbers, which are also infinite in extent. An extensional definition could be given in the following way, to take McGilvray’s example in Chomsky and McGilvray (2012):

Consider addition applied to a finite domain and range. For the natural numbers $\{1, 2, 3\}$ and no others, the function addition yields three ordered pairs with the first set of values the domain and the second, the range: $\langle\{1, 1\}, 2\rangle$, $\langle\{1,1,1\}, 3\rangle$, $\langle\{1, 2\}, 3\rangle$. There are no others.

Merge then is a function for which such an extensional listing of all the range sets cannot be stated. It is an *intensional* function: the range output can only be stated in terms of a description of the function’s operation on domain elements (its merging of them), rather

than in terms of the range output being an element from a given finite set of elements. Being a recursive function means that its range output can then be used as a domain input for further merge operations, which is what gives the function (and natural language generally) the property of discrete infinity. This kind of a function is needed to describe the internal biological substrate of natural language as we cannot do so by simply listing all the words and sentences that compose an individual's linguistic knowledge. We need to know what the operation is (Merge) and how it works, in order to then predict what linguistic knowledge an individual has. Also, importantly, a major source of evidence for UG has been precisely predicting what linguistic knowledge an individual will not have, what sentences are predicted to be ungrammatical, and hence judged unacceptable (something Chomsky's *Aspects* (1965, ch.1) is generally credited with firmly establishing).

2.4 Conclusion

Traditional philosophical views of the meaning of predication—in terms of property ascription, universal instantiation, set membership—are unsatisfactory. To the extent that they are substantively meaningful, they are perhaps best considered parts of a metaphysical scheme which only partially accounts for all the possible logical predications (see further section 7.3). Logic itself has definitions of predicates.¹⁶ Ultimately, for Frege a predicate is a function from objects to truth values (understood intensionally nowadays, so as a function from possible worlds to a function from objects to truth values). Functions aren't meaningful in themselves, though the elements of the domain and range sets (in so far as they reduce to objects) are. Determining, for any given sentence, what function(s) it contains appears to take us into the realm of information structure, the realm of comparison sets of sentences such that one can determine the fixed part (the function) and variable portion (the argument(s)) for any given sentence. Even identity sentences are ultimately analysed in terms of such functions for Frege: *The morning star is Venus* can be analysed as containing the $\langle e, t \rangle$ -type function: $\lambda x.[x \text{ is no other than Venus}]$ which *the morning star* could be applied to (equally, it could be analysed as containing the $\langle e, t \rangle$ -type function: $\lambda x.[\text{the morning star is no other than } x]$ which *Venus* could be applied to).

Functions have entered the landscape of syntax in the guise of theta roles, internal–

¹⁶Barry Smith (2005) has argued persuasively against what he calls 'Fantology' in the realm of metaphysics. His complaint is quite parallel to the way I consider resorting to logic as too reductive an approach to copular sentences. For Smith, resorting to the $F(a)$ of predicate logic to categorise reality exhaustively into the 'general' and the 'particular' is too reductive. For me resorting to predicate logic to categorise copular sentences exhaustively into 'predications' and 'identities' is too reductive. Jonathan Lowe's 'sortal logic' (originally (1980; 1982) but developed in numerous subsequent works) is consistent with addressing Smith's complaint.

external arguments, type-shifting operations (which ultimately map an expression of one type to its correlate of a different type, like Chierchia’s ‘nom’ operation, that shifts $\langle e, t \rangle$ -type *wise* to its nominal/objectual $\langle e \rangle$ -type correlate *wisdom*, *being wise*, *to be wise* understood as ‘the state of being wise’), and even the central combinatorial operation of UG itself (and possibly the only human- and language-specific aspect of the biological faculty of language according to Hauser, Chomsky, and Fitch (2002))—Merge. Attempts have been made to define predicates in terms of such functions, or combinations thereof (such as Williams’s ‘external theta role assigner’ definition of a predicate), though these have floundered in the face of empirical evidence.

So from what we have looked at in this chapter, it really isn’t obvious that there is a substantively meaningful notion of predication in logic or syntax which could be used to discriminate one class of copular sentences from another. In so far as compositional meaning is an effect of functional application—the definition of predication for Frege and his followers—we expect functional application to enter into any copular sentence, including identities (which is explicitly claimed by Frege, see section 2.2.7).

For those who wish to claim that there is an important distinction between types of copular sentence which is based on the logical distinction between predication and identity, they may consider there to be a refuge: the meaning of the copula itself, or rather of the linguistic constituents we find between the flankers in copular sentences, like *be* in English. They might try to argue that *be* is a meaningless piece of syntactic structure in predications—i.e. a real copula—but that it is a lexically contentful verb in identities. Chapters 4 and 5 will disagree with this move (and depart from Frege in this respect), arguing that *be* is never a lexically contentful verb, but—in copular sentences—always a copula. But note that the work of this chapter means that, whilst I agree that identities are distinct from other types of copular sentence, I don’t consider this a result of a logical relation which is mutually exclusive with a logical relation of predication which determines the meaning of all other copular sentences. Neither in the copula nor anywhere else in a sentence. I’ve tried to establish that already in this chapter, such that the work of chapters 4 and 5 should just be affirming why, in particular, it wouldn’t work to try and situate a logical distinction in the copula itself.

Ultimately, one main conclusion of this thesis will be that the effect of the logical distinction between predication and identity is orthogonal to natural language concerns. At least where predication is concerned, we’ve discovered that it can’t be used as a basis to distinguish copular sentences, if we think of the distinction between them as relating to their meanings. Predication isn’t a meaningful notion; ‘predicate’ as a generalisation doesn’t obviously have any kind of extension. This appears to be faithful to Frege, who introduced the

predicate (understood as a function) as part of the combinatorics of predicate logic, more so than part of the semantics.

Chapter 3

Can the philosophical notion of identity help describe or explain the taxonomy of copular sentences?

I shall not here discuss the difficult question whether “Tully is Cicero” exemplifies the classical uses of “Tully” and “Cicero” as names, or whether we should rather regard it as a proposition about these names in this use; that is, whether its analysis is something like “Tully is the same man as Cicero”, the names being used just as they might be in making historical statements, or rather something like “In history books the names ‘Tully’ and ‘Cicero’ are commonly used for the same man”.

(Geach, 1980a, 67)

3.1 Introduction

In the last chapter, we looked at the logical notion of predication, specifically the notion of predication that we find in Frege’s work, which has informed the foundations of contemporary natural language semantics. We did this because we wanted to know whether there was a substantive meaning of logical predication which could be used to distinguish those copular sentences deemed to be ‘predicational’ from those of the three other categories of Higgins’s taxonomy (specificational, identificational, and identity, see chapter 1). The premise of the argument here is that Higgins’s taxonomy is indeed a taxonomy of meanings, that it is the meanings (in any sense) of the sentences which distinguishes them (Higgins himself doesn’t use the word ‘meaning’, but rather ‘function’—in chapter 6, I analyse in more detail what

Higgins may have had in mind).

So we looked at logical predication in the last chapter and decided that it didn't have an obvious semantic extension in and of itself; it doesn't refer to an abstract notion of 'property ascription' (a subset of predicates may well do so: *red* in *the book is red* plausibly refers to the property of being red/redness; see section 7.3 for the view that such sentences are legitimate property ascriptions in metaphysical terms). But what about identity? A difficult matter to decide is whether the view 'predication is not semantically substantive in a way that can discriminate copular sentences' carries along with it the view that identity won't be either. The reason for this is that predication and identity are generally considered to be mutually exclusive in predicate logic.¹ It isn't clear to me how to proceed with this line of thought, so here I shall treat predication and identity independently, analysing them as distinct phenomena. Here, we turn to identity, especially as it's been studied in the analytic philosophy tradition. The aim, again, will be to see if this notion (or indeed these notions, as we'll see) of identity is what is informing the classification of copular sentences by linguists.

In the first part of the chapter—section 3.2—we consider whether there is a logical and/or metaphysical definition of identity that enters into copular sentences in natural language. For example, we might define identity in terms of the Identity of Indiscernibles part of Leibniz's law:

- (1) If, for any property, x has that property if and only if y has that property, then $x = y$.

Then my question is: does this give us a nice account of the meaning of the apparent identity relation in an identity statement like (2)?

- (2) Hesperus is Phosphorus

The first part of the chapter surveys philosophy for how it has sought to understand identity in a way that could be useful for understanding the meaning of copular sentences. In section 3.2.1 I look at the way that identity can be considered a univocal notion and what this means for the meaning of a sentence apparently expressing identity. In section 3.2.2, I look in more detail at Leibniz's law as a definition of identity to see whether (part of) it can be taken as somehow equivalent to the meaning of identity in an identity sentence.

The second part of the chapter—section 3.3—is concerned more directly with philosophical analyses of identity *statements* (generally considered to be part of philosophy of language). A basic copular sentence is the most typical example of an identity statement

¹Recall (from section 2.2.7) that Frege allowed for identity statements to be analysed as predications: $\lambda x.\lambda y.[y \text{ is no other than } x]$. Of course, if we turn that into $\lambda x.\lambda y.[y = x]$, then we can see that it reduces to logical identity ultimately in a way that is mutually exclusive with, say, $x \text{ loves } y$: $\lambda x.\lambda y.[Lxy]$.

that is given in the literature, but in principle an identity statement—in English—could also take such (more elaborate) forms as:

- (3) a. Hesperus is the same as Phosphorus.
 b. Hesperus is identical to Phosphorus.
 c. Hesperus and Phosphorus are the same/identical/exactly alike.
 d. Hesperus and Phosphorus are the same thing.

etc.² The most famous explicit reference to the meanings of identity statements is Frege's 'On Sense and Reference' (1892), but there are antecedents in Frege's own writing, as well as works that post-date this text that will be useful to survey. The question about these works will be the same as for the first part: Is there a meaning of identity statements from philosophy that we can compare with the equative copular sentences of natural language? The other crucial question is whether this meaning could be extended to any of the other classes of copular sentences, notably the specificational class, which as we saw in 2.1 have been argued to be reducible to identities.

Geach's epigraph to this chapter is a warning, but unfortunately I shan't be able to avoid 'the difficult question' about the meaning of these sentences that he raises; that is after all rather the point of this chapter, as I've just outlined. But if Geach found this question difficult then those of us with far inferior analytical powers should probably be wary of even trying. In fact, we needn't worry too much. We don't have to actually answer the question, we simply need to find some suggestions of answers to the question from philosophy and ask whether they are plausible accounts of the meanings of copular sentences that linguists have argued are 'identities'. Remember, crucially, that this includes specificational copular sentences for many linguists. The guiding intuition of this chapter is that the debate about whether specificationals are identities or not (see section 1.2.1) is ill-founded, at least partly because specificational meaning doesn't correspond to anything that has been claimed to be identity in the philosophical literature. Whatever else specificational sentences are doing, they are not stating identities, they don't mean that there is an identity between their two flankers. Also, note that the difficult question Geach poses may not exhaust the possibilities for the meaning of an identity statement; we shall compile a list of several meanings of identity statements in the conclusion.

²It is perhaps interesting that these are all copular sentences; English doesn't appear to have a lexical verb which states identity between its subject and object and if this is cross-linguistically true it would be evidence that identity is the sort of relation—if it is any sort of relation—that cannot be expressed with a lexical verbal, something which would be interesting in its own right.

3.2 ‘Pure’ definitions of identity from logic and metaphysics

It isn’t entirely obvious what field the study of identity in and of itself (as opposed to identity *statements* of natural language) belongs to. According to McGinn (2000), identity is a ‘logical property’; others (e.g. Lowe (2009)) think of it as a metaphysical relation. I don’t need to adjudicate that matter here. Whatever identity is in its ‘pure’ form, I think the least we could say is that it might enter into the meaning of an identity statement. Admittedly, this position should be defended in detail, but for now I’ll just take it as a working assumption i.e. in the same way that we might expect the action of kissing to somehow enter into the meaning of *kiss*, we might expect the ‘pure’ notion of identity to enter into the meaning of an equative copular, or the meaning of words like *same* and *identical*.

Frege (1892b, 56) gives a definition of identity which is quite commonplace. It is the unique relation that holds ‘of a thing to itself, and indeed one in which each thing stands to itself but to no other thing’ (‘that unique relation which, of necessity, each thing bears to itself and to no other thing’ (Lowe, 2009, 24)). It isn’t obvious that this is a definition of identity so much as a fact about identity. Equally, we can state facts about what kind of a relation it is: reflexive, symmetric, and transitive. In other words, for any x , x is identical to x is true (reflexivity), for any x and y , if x is identical to y then y is identical to x (symmetry), and for any x , y , and z , if x is identical to y and y is identical to z then x is identical to z (transitivity).

There are a few potential ways in which identity may be considered to fall into different ‘types’. I’ll sketch these ways briefly here but will conclude in line with what Lowe (2009, 71) calls ‘the classical view’, that ‘there is just *one*, univocal identity relation’ and not different types of identity. As McGinn (2000, 1) puts it there are not various forms of identity as there are shades of blue (in relation to the sentence ‘ x is blue’) or types of intelligence (in relation to the sentence ‘ x is more intelligent than y ’). Another way to put this is that saying something is identical with something else does not admit of the kind of ambiguity or under-determination of meaning we find in other lexical predications.

3.2.1 Putative types of identity

The central way in which identity has been classically sub-divided into apparently different forms, is the **numerical** (or token) **vs. qualitative** (or type) identity distinction. Numerical identity is an entity’s being that very same entity. For something to be numerically identical with my car it must be the actual object which is my car, not simply another car which is the

same make, colour etc. There can only be one object involved in the relation of numerical identity. In contrast, qualitative identity holds precisely when there is more than one object involved. Here, it is merely an object's qualities which are shared with another numerically distinct object. McGinn (2000, 2) says, 'It would be better to drop talk of 'numerical' and 'qualitative' identity altogether, speaking instead simply of identity and resemblance'. This seems appropriate especially when only *some* of an object's qualities are shared.³

As noted above, qualitative identity is really only resemblance, so not a type of identity. So there is no numerical identity relation between two distinct qualitatively identical entities. But what about the identity relation between the qualities of the entities, is that numerical identity? Well, if one believes that qualities are metaphysically real, e.g. if one believed in a theory of properties as ideal Platonic Forms—mind-externally independent entities with a real existence above and beyond the ordinary world of particular objects that we perceive by the senses (as set out in e.g. *The Republic* (Plato, c.380BC/1892))—then one could say that qualitative identity could be defined as distinct objects with numerically identical properties. The Daihatsu Terios-ness of my car would be that exact same numerically identical Daihatsu Terios Form which is a property of your car.⁴ This kind of stance towards the numerical–qualitative distinction is quite mainstream; as Lowe (2002, 24) puts it, most philosophers and logicians 'would say that when we speak of a thing becoming qualitatively different over time, we simply mean that this thing has numerically different qualities at numerically different times. . . On this view, all identity is really numerical identity, but we have to distinguish between the identity of a thing and the identity of its qualities'.

Two objects could in principle share every property i.e. for every property of x there could be a property of y which is numerically identical to it (based on universalism about properties, as just mentioned) and vice versa. Such a case would certainly be vanishingly rare on earth, but not theoretically impossible. In fact, there is a way of arguing that such a case would necessarily implicate numerical identity, according to 'Leibniz's law', that we'll discuss below in section 3.2.2. So really, we'd have to say that such cases are where two objects x and y share every property *except* the property of being identical to x (or being

³This usage is common in natural language; upon seeing a friend's new car for the first time, I may remark upon the fact that it is a Daihatsu Terios, like my car, by saying. 'I like your new car, it's the same as mine', even though it may be a different colour, year etc. Of course, it may be remarked from that example that the identity relation only actually holds between the *makes* of the two cars, such that what I said is really an elliptical version of 'It's the same *make* as mine'. And it may be further remarked that the identity relation is always of this sortally-relative kind. This is the thesis of 'relative identity' which I'll turn to briefly in this section, just below.

⁴Others, e.g. nominalists (at least ones like Quine (1964; 1981) who were nominalists about universals, if not about abstract objects), who don't believe in properties as metaphysically real entities independent of the objects which have them, clearly can't make this claim, as the properties of x (whatever they are) are necessarily not numerically identical with other distinct properties in a distinct object y .

identical to y); this will make sense after discussing Leibniz's law.

Another putative subdivision of types of identity really concerns identity *statements*, but I'm putting it in the first part of this chapter because it is generally understood as a debate about the metaphysics of identity, moreso than the meaning of identity statements. In essence, it asks whether sentences like *Hesperus is the same as Phosphorus* should be understood as shorthand for sentences which include a sortal term e.g. *Hesperus is the same **star** as Phosphorus*. A standard version of this view of **relative (vs. absolute) identity** allows that x may be the same F but not the same G as y .

But there is an immediate problem with saying this because 'the indiscernibility of identity' would be infringed and x and y would therefore not be numerically identical, so would only actually resemble each other. The indiscernibility of identity is the view that if x is identical to y , then every property of x is identical to every property of y . This is essentially the 'identity of indiscernibles'—that we saw earlier in the introduction, section 3.1 above—just with the inverse ordering of the flankers around the conditional; technically, Leibniz's law is either considered to be the identity of indiscernibles on its own, or the conjunction of it with the indiscernibility of identity, often by way of a single biconditional linking the flankers (Forrest, 2016). The reason why the indiscernibility of identicals is infringed in the case where x is the same F but not the same G as y is because x has the property 'being the same G as x ' (by the reflexivity of identity; I am assuming that relative identity is reflexive, like absolute identity) which y does not have, '[f]or if y has this property then it *is* the same G as x after all' (McGinn, 2000, 4), contradicting the starting premise that x could be the same F but not the same G as y in the relative identity view of things. So it seems that either x is not the same G as y , in which case x and y are simply distinct, or x is the same G as y , in which case x and y are absolutely identical, or at least, we cannot state that one has a property that the other doesn't.

Perhaps the issue about any alleged 'incompleteness'—as Geach (1972) put it—of identity sentences is rather related to the way that certain singular terms generally require sortal supplementation in order to, as McGinn (2000, 5) puts it, 'achieve determinate reference'. So if I say:

(4) This is the same as that.

then the demonstratives *this* and *that* are potentially indeterminate in reference in such a way that such an utterance may elicit a request for sortal supplementation, *the same what?* Equally, as McGinn (2000) points out, the request might be *What is the same as what?* To exemplify, imagine that I have two photos of the star Betelgeuse taken from different angles

(i.e. different positions on earth) or different times at night. And in both photos the rest of Orion's Belt (the constellation of which Betelgeuse is a member) is also visible. If I then utter (4), there is at least an ambiguity whereby *this* and *that* are either referring to a.) Betelgeuse, or b.) Orion's belt. In both cases I could resolve the ambiguity by saying:

(5) This is the same star as that.

(assuming it was obvious somehow which of the stars of Orion's Belt was intended). Or I could say

(6) This is the same constellation as that.

But equally, I could say:

(7) This star is the same as that star.

or

(8) This constellation is the same as that constellation.

This, then, is just trying to show that the perceived incompleteness of certain identity statements may be taken as a consequence of the indeterminacy of the referents of the flankers in them. When this reference is made more determinate, the need for sortal supplementation diminishes. Sortal supplementation, then, is not somehow a necessary aspect of the identity relation, moreso than a linguistic way of helping to determine the reference of two flankers.

3.2.1.1 Conclusion about putative types of identity

The conclusion of all this is that whatever identity is, it looks like (and is quite standardly accepted to be) a univocal notion.⁵ The apparent distinctions we've looked at either confuse identity for resemblance (numerical vs. qualitative), or relate to the often indeterminate reference of the terms of an identity sentence (absolute vs. relative).

Now, in terms of point of this chapter and this thesis, what do these conclusions mean? The lesson is that the 'pure' notion of identity isn't actually going to take us very far in understanding identity statements. Why not? Because we have established that identity,

⁵I have overlooked another distinction that is often brought in to debates about whether identity is univocal or not: **necessary vs. contingent identity** e.g. the difference between *Cicero is Tully*, where Cicero could not have *not* been Tully, and *Trump is the 45th president of the USA* where he may well not have been. I think this is quite generally taken to be related to the modal properties of identity *sentences* rather than of identity understood in a pure logical or metaphysical way, and so not a threat to the thesis of the univocality of identity, see McGinn (2000, ch.1). This interpretive difference between necessary and contingent identity statements (which the reader may notice aligns with the difference between an equative and specificational copular sentence), is something Kripke (1980) notably discussed.

properly speaking, is just numerical identity, which is just the fact of one object being the object it is and no other object (or one property being the property it is and no other property etc.). Strictly speaking, the fact of an object being the object it is isn't something one *says about* an object. In fact, the only putatively informative utterance the univocal notion of numerical identity gives rise to—for any given object *o*—is to state the numerical identity of *o*: *o* is the object it is, and no other object. But it is immediately clear that this is not what identity statements say, indeed this is not what anyone says outside of philosophical discussions about the nature of identity. If we assume that identity statements are a common part of informal discourse, then we have to conclude that they are not simply statements about the numerical identity of an object. The foregoing has therefore served to lead to the conclusion that the mainstream analysis of identity as a logical or metaphysical property—as interesting as it is in its own right—does not enter into an understanding of identity statements. Another way of stating this conclusion, I think, is to say that the pure notion of identity only gives rise to statements which are tautologies: the numerical identity of *o* with itself gives rise to the sentence:

- (9) a. *o* is (the same as/identical to) *o*.

Again, in so far as identity statements are a common part of informal discourse, and are seen as informative in some way, then it is unlikely that the 'pure' definition of identity which only results in uttering tautologies captures their meaning.

Wiggins (1965, 48) makes this point in a slightly different way:

Now there is no absurdity in a relation which holds between a thing and itself. A man may kick, feed, make provision for, &c., himself. The trouble with self-identity is that (unlike these reflexive transactions between part and whole, whole and environment) it is true of any individual whatever. If 'self-identical' marks nothing off, how can it be a predicate or ascribe a genuine property?

Wiggins's conclusion seems to be the same as the one I have come to here: For Wiggins, identity (understood in the 'pure' way) cannot be a true predicate. I agree, in so far as we are talking about natural language. As I've said, I think the standard analysis of 'pure' identity gives us a notion which—as logically/metaphysically substantive as it may be—doesn't obviously translate into a notion that would ever be asserted in natural language (except to define the logical/metaphysical notion of identity).

Having said this, we shall continue with the 'pure' analysis of identity a little longer, looking specifically at the definition of identity given in Leibniz's law.

3.2.2 Leibniz's law

We've already made passing reference to Leibniz's law in this chapter, but it deserves to be focused on a little more to see whether it could help us understand the meaning of identity statements.

A definition of identity may be given in terms of Leibniz's law (P a property):

$$(10) \quad x = y \text{ iff } \forall P, Px \leftrightarrow Py$$

This is the biconditional I mentioned earlier (whilst discussing absolute–relative distinction in section 3.2.1) which combines the indiscernibility of identicals (going left to right with a straight 'if–then' conditional) with the identity of indiscernibles (if you inverted the flankers and had a straight 'if–then' conditional). The *Principia Mathematica* (Whitehead & Russell, 1963, vol. I, definition 13.01) gives a definition of identity which is essentially Leibniz's law relative to a notion of 'predicative functions':

x and y are to be called identical when every predicative function satisfied by x is also satisfied by y .

In fact, McGinn (2000, 8) says that Leibniz's law 'is the only plausible attempt to define identity'. So a condition, and the only condition, on x and y 's being identical is that they are indiscernible, i.e. that they have all and only the same properties, and therefore could not be differentiated by 'discerning' a property that was in one but not the other. Now, we might immediately respond rather as we did earlier when discussing qualitative identity: whilst it may be an exceedingly rare event on earth to find two things which are truly identical with respect to all their properties—e.g. identical twins with not only exactly the same physical properties (down to e.g. the structure of each cell in their body) but also identical knowledge of French and all other subjects (considered as properties) etc.—this is just an empirical fact about life on earth rather than some kind of theoretical impossibility, and so we can apparently allow for total qualitative identity of distinct objects without it necessarily determining numerical identity. So the way to make Leibniz's law a better definition of identity is to say, as Black (1952, 153) does:

If two things, a and b , are given, the first has the property of being identical with a . Now b cannot have this property, for else b would be a , and we should have only one thing, not two as assumed. Hence a has at least one property, which b does not have, that is to say the property of being identical with a .

(This is essentially the same reasoning behind the way I argued that relative identity is not a *type* of identity in section 3.2.1 above.) So any distinct things will always have at least

one property that is not shared. From this then, if we stipulate that a and b do not have any properties that are not shared, i.e. if all and only the properties of a are the same as the properties of b , as Leibniz's law states, then we can infer that there is in fact only one thing i.e. that $a = b$.

Frege (1884, 76) quotes Leibniz's own rendering of Leibniz's law:

Things are the same as each other, of which one can be substituted for the other without loss of truth. [Frege's translation of: *Eadem sunt, quorum unum potest substitui alteri salva veritate.* (Leibniz, 1840, 94)]

This can be viewed as equivalent to the law stated in terms of properties just above. Presumably the idea here is that a thing having a property ascribed to it is the way in which there will be truth or not. So where one thing can be substituted for another without loss of truth, we can understand that thing b , say, is substituted for a , and that every property that is ascribed to thing b results in the same true or false property ascription as we find for thing a . And this would only be possible if things a and b shared all and only the same properties. And in that case, according to Leibniz, they are the same. So thinking in terms of total sharing of properties or in terms of substitution without loss of truth are equivalent ways of thinking about Leibniz's law.

There are certainly difficulties with Leibniz's law. For a start, it is circular. To take the last version we just gave, the substitution of one thing for another must be understood as meaning substitution in the *same context* e.g. the same sentence, a notion that requires an understanding of identity in the first place. Wiggins (1965, 49) makes the circularity point slightly differently:⁶

[R]elations have to have terms to hold between and hence presuppose the identification of these. So surely they presuppose the identity of these. How then could identity itself (or, its negation, diversity) be a relation?... Predication presupposes identification. Identification presupposes the possibility of re-identification. This last presupposes the possibility of identity-statements.

⁶This is a challenge specifically to the notion of identity understood as a *relation/predication*. It is, after all, an unusual relation, the relation that only ever holds between an object and itself and is the only such relation. The issue of whether identity is a relation comes to the fore when we are dealing with identity statements. As Russell (1903, p.63) says,

For, it may be said, identity cannot be a relation, since, where it is truly asserted, we have only one term, whereas two terms are required for a relation.

Frege explicitly tries to overcome this problem by precisely locating a *difference* between the two sides of a sentence stating an identity, such as *Hesperus is Phosphorus* or *Hesperus is identical to Phosphorus*; this begins in section 3.3.1.

An influential criticism of Leibniz's law stems from Black (1952), who contemplates a denuded universe with nothing but two exactly similar spheres, symmetrically placed, such that 'every qualitative and relational characteristic of the one would also be a property of the other' (Black, 1952, 156): so indiscernible spheres, yet clearly there are two of them. Above, we showed the real interest of Leibniz's law by making the point that an object *a* will always have a property *being identical with a* which another object *b* could not have, even though all other properties of *a* and *b* are shared. How does Black deal with this point, for surely we could say the same thing here: sphere *a* has the property of *being identical to sphere a*; sphere *b* either has this property—in which case it is in fact sphere *a*—or it doesn't, in which case it does not in fact share all its properties with sphere *a*, and is simply a distinct object. One way Black gets round this is just to say that it is impossible to name the objects in the first place, or rather that such a naming would require the presence of some third element (an observer or namer), which would then create a universe which is not the hypothetical one under consideration. Without names, the crucial properties that strengthen Leibniz's law as a definition of identity are not formulable. As Black (1952, 163) concludes then, '... it is logically possible for two things to have all their properties in common', or at least to invent scenarios in which Leibniz's law doesn't seem to differentiate two objects.

3.2.2.1 Can Leibniz's law serve as a meaning of an identity sentence?

Beyond these criticisms, could Leibniz's law—as a potential definition of identity—enter into the meaning of an identity sentence? In other words, does *Hesperus is Phosphorus* mean 'For any property *P*, if Hesperus has that property then Phosphorus has that property'? Perhaps we could render it into idiomatic English as:

- (11) Hesperus has all and only the properties that Phosphorus has. *or*
 Hesperus and Phosphorus have all their properties in common. *or*
 Hesperus and Phosphorus are exactly alike.

Let's try to put these statements into a discourse context:

- (12) (Ancient Greece c.530BCE)
 Anaximenes: Hey Pythagoras, I got a great view of Hesperus last night. It really is my favourite star. I much prefer it to Phosphorus.
 Pythagoras: I actually just worked out that Hesperus *is* Phosphorus.
 Pythagoras': I actually just worked out that Hesperus and Phosphorus are exactly alike/share all their properties in common.
 Pythagoras'': #I actually just worked out that for any property *P*, if Hesperus has

that property then Phosphorus has that property too.

The oddity of Pythagoras'' is perhaps just the oddity of talking in logic-language in an otherwise idiomatic discourse. Pythagoras' doesn't, on the face of it, seem pragmatically odd. But, crucially, we are aiming at a meaning equivalence to the copular sentence answer of Pythagoras, and it isn't obvious to me that Pythagoras' is such an equivalence. Most people would take it to mean that Pythagoras is saying that the two distinct objects Hesperus and Phosphorus exactly resemble each other. People versed in Leibniz's law (as we are here by now) may believe that for two objects to be exactly alike means that there aren't in fact two objects, for if Hesperus has the property of 'being identical to Hesperus' then either Phosphorus has it, in which case it is not a distinct object to Hesperus, or it doesn't, in which case it is a distinct object to Hesperus. So knowing that Hesperus and Phosphorus are exactly alike—understood in this 'strict' way, we might say, of exact likeness—allows us to deduce that Hesperus is identical to Phosphorus. But this, I think, shows the way that whatever the ' $\forall P, Px \leftrightarrow Py$ ' side of Leibniz's law is taken as, it doesn't obviously fulfil the same discourse function as *Hesperus is Phosphorus* in the context of (12) (which seems like a standard equative copular context).⁷ For example, it wouldn't obviously be tautologous for Anaximenes to continue the conversation in (12) in either of the following ways (which seem essentially equivalent):

(13) Pythagoras': I actually just worked out that Hesperus and Phosphorus are exactly alike.

Anaximenes: I see, in which case they must be the same object.

Anaximenes': I see, in which case Hesperus *is* Phosphorus.

It certainly would be tautologous for Anaximenes to continue the first Pythagoras reply in (12):

(14) Pythagoras: I actually just worked out that Hesperus *is* Phosphorus.

Anaximenes: #I see, in which case they must be the same object.

Anaximenes': #I see, in which case Hesperus *is* Phosphorus.

Equally of course (and as the biconditional of Leibniz's law allows), the deduction could go the other way:

(15) Pythagoras: I actually just worked out that Hesperus *is* Phosphorus.

⁷For now, we can understand 'discourse function' informally as just whatever conversational move an interlocutor appears to be making with his/her utterance. Later, in chapter 6, I shall try to clarify the notion of a function of a sentence.

Anaximenes: I see, in which case they must be exactly alike (in the sense of Leibniz's law).

So, to re-iterate, the fact that (14) is a tautology where (13) and (15) are not is intended to show that whatever else we can say about the equivalence of the meaning of *Hesperus is Phosphorus* and *Hesperus and Phosphorus are exactly alike*, they don't appear to play the same discourse function.

Let's now apply these Leibniz's law-based paraphrases to a specificational copular sentence, as these have been argued (by e.g Heycock and Kroch (1999), Mikkelsen (2005), and Heller (2005)) to be identities ultimately:

- (16) The culprit is Brian.
 =
 The culprit and Brian are exactly alike/have all their properties in common, *or*
 For any property P , if the culprit has that property then Brian has that property.

I think these paraphrases would, at best, be an implication of the sentence, some fact about the world that one could deduce from *The culprit is Brian*, but it wouldn't obviously be the meaning (particularly not in the sense of 'discourse function') of the sentence. In the specificational sentence's most obvious context, it would be bizarre:

- (17) At the scene of a crime
 Detective: Sergeant, do we know who is the culprit of this heinous act?
 Sergeant: #Yes detective, the culprit and Brian are exactly alike, *or*
 Sergeant': #Yes detective, for any property P , if the culprit has that property then Brian has that property too.

The oddity of (17), I think, is to do with the way that the function of a specificational sentence—i.e. specifying a value for a variable, or giving a one item list in Higgins's (1979) analysis—is not obviously fulfilled by the Sergeant's answer (even in the idiomatic version). There is a clear Question-Under-Discussion (QUD) (Benz & Jasinskaja, 2017) in (16) whereby the Detective is requesting the Sergeant to complete a list headed by the heading 'the culprit'. The Sergeant's answer, for a start, brings in a second entity—Brian—and then appears to say something about the relation between Brian and the culprit. But Brian was no part of the QUD. In (12), on the other hand, both Hesperus and Phosphorus were at least part of the 'common ground' of the discourse context. In any case, it seems fairly clear that neither equative copular sentences nor specificational copular sentences express identity

in exactly the same way as Leibniz's law.⁸

3.2.3 Eliminating identity and identity as undefinable

The last view about 'pure' identity that we'll touch on here is Wittgenstein's view in the *Tractatus* and views about identity being undefinable. Wittgenstein was clearly arguing that the identity sign should be eliminated from logic, so his thoughts more clearly relate to logical identity than previous sections, which were more to do with identity as a metaphysical notion. If identity can indeed be eliminated from logic, then it would of course be impossible to reduce an identity *statement* to a representation of logical identity. So let's look at what Wittgenstein said.

[§5.5301] It is self-evident that identity is not a relation between objects... [§5.5303]
Roughly speaking, to say of two things that they are identical is nonsense, and to say of one thing that it is identical with itself is to say nothing at all. (Wittgenstein, 1922)

Wittgenstein then proposed a notation which eliminates '='—the sign for identity—in favour of capturing the same information by 'sameness of sign'. So, for example, Wittgenstein wouldn't write (5.532)⁹:

$$(18) \quad (\exists x, y). f(x, y). x = y$$

but rather

$$(19) \quad \exists x. f(x, x)$$

Interestingly, despite the fact that most philosophers disagree with Wittgenstein's elimination of the identity sign, they do resort to the same point Wittgenstein makes, but to achieve a different end. For example, McGinn (2000, 8) argues that identity is presupposed in the notion of a variable in order to say that identity cannot be defined non-circularly:

⁸McGinn (2000, 8) makes a related point: he says that *Hesperus is Phosphorus* is a 'singular relational statement about a given object' where the ' $\forall P, Px \leftrightarrow Py$ ' side of Leibniz's law involves, 'quantifying second-order-wise over properties and making a statement of generality' turning *is* from a two-place relation word into a quantifier and a biconditional which expresses 'a higher order condition on properties'. This seems like a fair point: logical forms of natural language sentences are taken to express their meanings. Therefore if a natural language copular sentence apparently expresses a relation as part of its meaning then it is rather odd that its logical form expresses a non-relational universally quantified biconditional statement. And the infelicity of such statements as equivalents for the identity sentences of which they are the definitions in natural language contexts (such as the examples in (17) and (12)) is consistent with this point, I think.

⁹The second dot is a conjunction sign.

[I]f we say ‘for some x , x is F and x is G ’ we are making tacit appeal to the idea of identity in using ‘ x ’ twice here: it has to be the same object that is both F and G for this formula to come out true.

But for Wittgenstein, the fact of a subsequent use of a variable x presupposing its identity with the first use of variable x is a reason to eliminate the identity sign. We already have identity implicit in the use of a bound variable, so let us just repeat a given variable rather than resorting to two different variables and identifying them as the same as each other, he claims. A possible response to this would be to say that the notion of *distinctness* is perhaps *not* necessarily implicit in the use of *different* variables. But this doesn’t seem like a good response; it seems clear that in Wittgenstein’s system distinctness *is* implicit in the use of different variables. He would probably just say that this falls out directly from sameness of variable being the only possible sameness; non-sameness of variable would mean necessarily non-same reference of the variable.

As I said above, Wittgenstein (in the *Tractatus*) is concerned with logic, not natural language. However, could we claim that ‘ $\exists x.f(x, x)$ ’ could serve as a representation of *Hesperus is Hesperus*? I don’t think so. In §5.534 Wittgenstein calls ‘ $a = a$ ’ a ‘pseudo-proposition’ that ‘cannot even be written down’ in his system. If ‘ $\exists x.f(x, x)$ ’ were to serve as a natural language predicate, it would plausibly be something like *wash* in *John washed himself*. So for Wittgenstein, neither informative identity statements like *Hesperus is Phosphorus* nor non-informative identity statements like *Hesperus is Hesperus* are representable in his system. It’s certainly interesting to wonder why logic may not need informative identities where natural language does. The obvious reason that springs to mind is that there is no ambiguity about the denotations of constants and variables (under an assignment) in logic, given the interpretation function (see section 2.2.8), whereas ambiguity of reference is widespread in natural language.

Wittgenstein’s elimination of identity from his logical language shouldn’t be confused with the metaphysical position that identity is undefinable, even though we just saw that a similar argument is made to draw both conclusions. The view that identity is not definable is quite mainstream, e.g. Lowe (2009, 25): ‘Identity may simply be one of those conceptual primitives—like, perhaps, existence—of which it is true to say that if one did not already grasp what it is, it could not be explained to one’. But what would that mean for identity sentences? I don’t think it would mean that they are meaningless. For a start, McGinn (2000)—which is ‘an old-fashioned and sensible view of identity’ (Blanchette, 2002)—is keen to maintain that identity is a *real* relation, despite its undefinability. This, then, leads to the position that identity is on the one hand absent (Wittgenstein) or at best undefinable (McGinn and the mainstream) as a property of a logical language or as a metaphysical

notion, but that identity sentences of natural language are perfectly meaningful, just not with a meaning that corresponds with logical/metaphysical identity (due to its either being non-existent or indefinable). This would make the meaning of *Hesperus is Phosphorus* an entirely linguistic affair, irreducible to logic or metaphysics.

Russell originally agreed with Wittgenstein, but later changed his mind (Russell, 1959, 85–6). Essentially, Russell felt that Wittgenstein had overlooked the full force of Leibniz’s law. Leibniz’s law was the definition of identity that Wittgenstein was reacting against in the *Tractatus* (specifically Whitehead and Russell’s version of it in *Principia Mathematica*); Wittgenstein said this ‘won’t do’ because, ‘according to it one cannot say that two objects have all their properties in common (Even if this proposition is never true, it is nevertheless *significant*)’ (Wittgenstein, 1922, §5.5302).¹⁰ But this statement, for Russell, is what gives away that Wittgenstein hasn’t appreciated the full force of Leibniz’s law. As we saw in 3.2.2, the crucial part of the argument is that a will have the property of being identical to a , which b either has—in which case b is a —or which it doesn’t, in which case b is not a . As Russell puts it (1959, 86): ‘[I]f two things have all their properties in common, they cannot be *counted* as two, since this involves distinguishing them and thereby conferring different properties upon them’. So it is impossible for two distinct objects to have all their properties in common, as Wittgenstein seems to believe in §5.5302.

Generally modern logicians appear to have rejected Wittgenstein’s view and consider that a first-order logical language (with or—like a Quinean first-order language—without singular terms) without an expression for identity is strictly less expressively powerful than one with an expression for identity (Jonathan Lowe, p.c.). Wittgenstein’s view, then, would be a final nail in the coffin for the idea that logical identity (or any kind of ‘pure’ notion of identity from philosophy) can be contributing to the meaning of an identity statement. However, its controversial status means that we shouldn’t rely on it for this position. The conclusions that we drew in 3.2.1.1 and about Leibniz’s law in 3.2.2.1 will have to be our strongest arguments to conclude that explorations of identity in logic and metaphysics do not obviously lend themselves to an explication of the meaning of identity statements.

3.3 Meanings of identity statements

Having looked at the ‘pure’ definitions of identity, we shall now turn to see how writers have proposed to understand the meanings of identity statements. It perhaps isn’t immediately obvious that these are separate things, so I’ll say a few words in this pre-ambule to justify the

¹⁰This is Russell’s version. The Pears and McGuinness translation (1922) has it as: ‘Even if this proposition is never correct, it still has *sense*’.

separate parts of this chapter.

Let's take the classic identity sentence:

(20) Hesperus is Phosphorus.

Intuitively, it states an identity between Hesperus and Phosphorus. Then we might want to know what the definition of identity is, in such a way as to understand what it is that is being stated to hold between Hesperus and Phosphorus. And so we will look at definitions of identity in order to help us understand (20). And that's what I did in the first part of this chapter; there were two main definitions: 1.) Numerical identity, which is the unique relation that an object has to itself and nothing else, 2.) Leibniz's law that if two things are indiscernible they are identical and vice versa. For the second of these in particular, we looked at how it might serve as a meaning of an equative copular (in section 3.2.2.1). After all, our central interest here is in determining if there is a notion from philosophy which could be used to classify one or more types of copular sentence. Where Leibniz's law is concerned, we concluded that it wasn't *prima facie* unacceptable to utter an idiomatic equivalent of it in place of an equative copular (example (12)), but it wasn't obviously performing the same discourse function, something I take to be an aspect of sentence meaning that could be used to distinguish classes of sentence.

So, as I stated in the outset to this chapter, there may be something to be taken from such pure enquiries into the nature of identity. It seems fair to say that identity can be studied in its own right, as a logical or metaphysical property. The first definition in the paragraph immediately above (numerical identity) looks like a statement of pure metaphysics, in so far as objects and relations are metaphysical notions ultimately (though the e.g. necessary reflexivity of the relation which is stated in the first definition may be generally considered more of a 'logical property' of a relation rather than a metaphysical fact about it). By stating what the identity relation is, we needn't produce an identity statement, and most metaphysicians would presumably consider that the relation would persist in a universe which did not contain anything stating identities. An object would still be identical with itself, regardless of whether this could be represented by anyone/thing else, they would maintain. Numerical identity, then, is a metaphysical relation holding of objects, and is oblivious to human interests.

Equally, however, as pointed out above (section 3.2.1.1) it may be argued that humans aren't very interested in numerical identity. That some object is identical to itself is certainly never informative, and statements of such metaphysical relations like

(21) Hesperus is Hesperus.

are tautologies. Whatever the connection is, then, between ‘pure’ identity and *informative* identity sentences of natural language, it isn’t obvious that informative identity statements are simply straight representations of this metaphysical relation, the way that, for example:

(22) The object Fido instantiates the kind dog.

may be considered to be a straight description of the metaphysical relation of instantiation holding between the object Fido and the kind dog (to use the ontological scheme of Lowe (2009) as an example). We just don’t seem to ever need to state that numerical identity holds of any particular entity, in the same way as we don’t tend to state that a given object has mass, a given person has a self, or other examples where we wouldn’t ever be telling anyone anything they don’t already know (or, rather, know in virtue of knowing what objects, people etc. are).¹¹

So we’ve seen that metaphysics gives us a notion of identity which may only rarely be stated (in philosophy classrooms, for example), because stating it amounts to a tautology.¹²

¹¹In that respect it seems like an analytic truth. Of course, if Frege was right to consider that a priori knowledge (e.g. knowledge of mathematics) could be reduced to analytic truths, then there are certainly informative statements that might in principle be reduced to analytic truths. In Frege’s logicist project, this meant translating mathematical statements into synonymous statements of his *Begriffsschrift* logical language; as Gödel (1990, 119) put it: “Frege was chiefly interested in the analysis of thought and used his calculus in the first place for deriving arithmetic from pure logic”. To take the example of ‘recollecting’ a priori knowledge Plato gives in *Meno* (Cottingham, 2008, 3–12): to learn that a line running diagonally between two opposite corners of a square can be used to create a new square which is exactly double the area of the original square may be considered informative to someone who needs to know how long to make the sides of a square which is double the area of another given square. If logicism is correct, this sort of mathematical knowledge, ‘discoverable by the mere operation of thought’ (Hume, 1748, 24), could be reduced to analytic truths.

¹²This seems to be consistent with Jespersen’s remarks related to identity statements:

But as we have seen [I am unable to find a discussion of this earlier in *Philosophy of Grammar*, AW], perfect identity is rare, and it is important to remark that the linguistic “copula” *is* does not mean or imply identity, but subsumption in the sense of the old Aristotelian logic, which is thus in closer accordance with grammar than the so-called logic of identity (Leibniz, Jevons, Høffding). According to the latter the sentence “Peter is stupid” should be analyzed as “Peter is stupid Peter,” or, as it is also maintained that the substance of the predicate influences that of the subject, we obtain perfect identity only by saying “Stupid Peter is stupid Peter.” In this way, however, the character of communication from speaker to hearer is lost; by the words “is stupid Peter” the hearer is told nothing more than he had heard at the beginning, and the sentence has no value whatever. Ordinary mortals, therefore, will always prefer the formula “Peter is stupid,” by which Peter is ranged among those beings (and things) that can be called “stupid.” (Jespersen, 1924, 154)

Jespersen’s remarks about the meaning of the copula are interesting in their own right, and so is what he says about ‘logic of identity’ (which was an immediate precursor to Frege’s logic) not being in close accordance with grammar, consistent with my view in this chapter. But the way he refers to ‘perfect identity’ here leads me to believe that he also considered it to always be essentially the statement of tautologies. Of course, informative identity statements do have communicative value and are not tautologies, therefore Jespersen would presumably not have considered them ‘perfect identities’, though I’ve unfortunately failed to find an

Given that identity statements are generally not tautologies, it seems that we are doing more than merely describing a metaphysical relation. Or so it seems. Are we even in fact stating the metaphysical relation at all? If not, then what is the relation between the metaphysical relation of identity and the linguistic meaning of an identity statement? I think the time has come now where we will simply have to look at some proposed meanings of identity statements to help us answer this question.

So in this section, we'll move on to analyses which are more clearly dealing with identity *statements*, more so than definitions that were dealing with identity as a logical or metaphysical relation. The aim here is the same as the first section (and the previous chapter, which dealt with predication). We want to see whether there is anything that we can take from this and bring to understanding the copular sentence taxonomy as a taxonomy of 'meaning types'. So we want to know if there are any meanings of identity statements from the history of analytic philosophy which we might be able to bring in to our analysis of equative copulars and specificational copulars, both of which have been treated as logical identities in linguistics. This means that our focus isn't so much on the questions that analyses of identity sentences were traditionally used to address. For instance, 'Frege's puzzle about identity statements' (Zalta, 2017, §3.1), set out in his 'On Sense and Reference', is centrally concerned with the semantics of proper names, with the problem of how $a = a$ can mean something different from $a = b$ (Frege, 1892b) given a direct reference view of proper names. The set-up to the puzzle doesn't actually have to make any reference to the meaning of identity statements (in terms of the meaning of the identity relation that they apparently state); it simply acknowledges the intuitive fact that there is some kind of difference in meaning between, say, *Hesperus is Hesperus* (a tautology) and *Hesperus is Phosphorus* (not a tautology). So we need to go in to these texts and try to just pick out any explicit mentions of the meaning of the identity relation we can find.

3.3.1 Frege's '2-signs' theory of identity statements

Dummett (1981, 279) is a nice link between Frege's view of identity in a pure way (which Dummett refers to by talking about what the identity sign '=' stands for below) and the meaning of an identity statement, and is consistent with what we said immediately above and in section 3.2.1.1:

The sign of identity is taken by Frege as standing for a relation which everything has to itself and nothing has to anything else. The informational content of a

explicit statement of what he would have considered their meanings. Elsewhere, in Jespersen (1933), he gives the example *So that's that!* as a 'perfect identity' which is a 'rare case'. But I think this expression has a more idiomatic meaning than mere tautology.

statement of identity can therefore never amount to a particular object's having the relation of identity to itself, for this is never news. Rather, it is just in virtue of the truth-conditions for statements of identity that we may learn from them that two names with different senses have the same object as their referent.

The notion of *senses* relates to Frege's 'On Sense and Reference' view of identity statements that we'll look at later (section 3.3.2).

Before 'Sense and Reference' though, Frege outlined a view of identity statements in section eight of the *Begriffsschrift*, which is cited as the source of what is sometimes called (by e.g. Wiggins (1965)) Frege's '2-signs theory', which relates to the way that identity statements can be viewed as statements where there are two signs for the same 'content' (to use Frege's own term). In this section, we shall look at the theory as Frege presents in the *Begriffsschrift*. Initially, however, we need to clarify an issue which is sometimes overlooked in discussions of identity statements.

3.3.1.1 The impossibility of identity statements merely applying between expressions

In the *Begriffsschrift*, Frege denies identity statements apply 'only to *expression*, not to thought; as though we had no need of two symbols for the same content' (Frege, 1879, 11); so, whilst the *Begriffsschrift* theory is called a '2-signs' theory, it needs to be stressed that identity statements are not *merely* statements that pertain to names/signs for Frege. If they did, as Frege says, we wouldn't really need two names for the same object.

This is a subtle but quite important point as a naive view of identity statements might indeed claim that they were merely stating that there were just two labels for one object. So I just want to spend this section trying to clarify why that isn't a coherent position in the first place.

The first thing to say is that names, labels, 'expressions' generally (to use Frege's term), are always *of* something. The notion of a label which is not a label for anything seems self-contradictory. So any given label will always have some thought contents or entity in the world associated with it, and in that case an identity sentence will always state an identity between that label *with its 'thought'* (to use Frege's term) and something else. Note that we could, just about, refer to a label using the label itself, in an identity sentence. In that case we would 'mention' the label, rather than 'use' it (Quine, 1940, 23–7). In writing at least this seems possible with the use of quotation marks:

(23) 'John' is 'George'.

But of course, any such identity statement where two different labels are mentioned will always be false. The name ‘John’ with four letters is clearly not the same label as the label ‘George’ with six letters.

One other thing to say about identity sentences not applying *only* to expression: it doesn’t seem rational for there to be two labels for an object if there was no *need* (or just no reason at all) for more than one label. The fact, therefore, of there ever being more than one label for something must reflect a reason for it, and this will serve to distinguish the labels beyond the mere *expression* difference of them being different labels (i.e. composed of different letters, number of letters, pronunciation etc.). Naming isn’t so arbitrary that one object is given two names for no reason. Every ‘natural’ instance of one object having two names has a rationale for the multiple names: *Hesperus* is a star viewed in the evening, whereas *Phosphorus* was a star seen in the morning; *Tully* is an anglicization of the name of the Roman general, Marcus Tullius Cicero—*Cicero* as he would have been known in the Rome of his own time; *George Orwell* is the pen-name of the person whose birth name was *Eric Blair*; *Clark Kent* is the earthly alias of the Kryptonite superhero *Superman*. In all these cases, there’s a reason for the duality of names, contexts where one of the names is used rather than the other. And, of course, the users of one of the names may well be unaware that the name-bearer also has another name. A name-bearer may have two names because two communities of speakers have, in separate times and places, given her a different name. But this is still a reason for a name-bearer having two names, and an identity sentence between them would be relating both the label and the thought ‘which is what you call her’ with ‘which is what we call her’. The point is that names are never really simply interchangeable in such a way that a relation between them would be between a name only.

To re-iterate, the examples above give us examples of what Frege called (in the *Begriffsschrift*) a ‘thought’ attached to the names. The thought attached to *Hesperus* is ‘a star viewed in the evening’; the thought attached to *Clark Kent* is ‘the earthly alias of *Superman*’. This ‘thought’ shouldn’t be confused with the ‘content’ of these expressions, which in both cases is an object in the world. So, really, there are three levels of description in Frege’s 2-signs theory: names, thoughts, and contents.

3.3.1.2 Articulating Frege’s 2-signs theory

Getting back to the *Begriffsschrift*, despite pointing out that identity statements don’t merely apply to expression, Frege does however say that they do ‘relate’ to names rather than contents, and this he says is unlike negation and conditionality which ‘relate to contents’, and generally it seems:

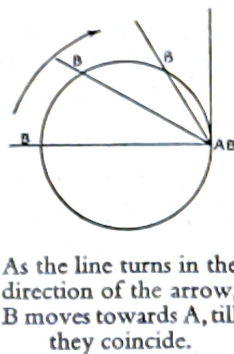


Figure 3.1: Illustration of Frege's example, Frege (1892b)

Elsewhere, signs are mere proxies for their content, and thus any phrase they occur in just expresses a relation between their various contents; but names at once appear *in propria persona* so soon as they are joined together by the symbol for equality of content; for this signifies the circumstance of two names' having the same content. (Frege, 1879)

This then is the crux of Frege's *Begriffsschrift* theory: two names having the same content. Identity statements are at once a relation between names; as the last section showed, names always carry with them a 'thought' associated with them. In the *Begriffsschrift*, Frege calls this 'the way of determining' a referent. He gives an example to clarify his view:

Take a point on the circumference of a circle, and call it point *A*. What have we done? We have determined a point from its being 'directly given in experience'; that's the first thing. Secondly, we have named it '*A*'. Now, imagine that there is a straight line running through the circle at point *A*. Then, let's determine another point in the following way: the point of intersection between this line and the circumference of the circle. Let's call this point *B*. In relation to this point, point *A* could be the pivot of the rotation of the line. So if, indeed, we rotate the line, what happens is that point *B* travels around the circumference of the circle. So, again, we have done two things: firstly, we determined a point; this time, we determined it in a mathematical way (roughly, the point on a straight line pivoting around point *A* which lies on the intersection of the line with the circumference of the circle), then we named it point *B*. Now, when the straight line runs through the diameter of the circle, points *A* and *B* are on opposite sides of the circle. However, when the straight line runs perpendicular to the diameter, points *A* and *B* coincide. The illustration in figure 3.1 makes this much clearer.

So, the need for the identity sign arises in order to to say that these different names ('*A*' and '*B*') corresponding to different ways of determining a point (*A*: direct experience, *B*:

the point corresponding to the line's being perpendicular to the diameter), share the same 'contents' (to use Frege's word in the *Begriffsschrift*; this can be correlated with 'referent' or 'denotation' here). In Frege's (1879, 11)'s exact words:

The same content can be fully determined in different ways; and *that*, in a particular case, *the same* content actually is given by *two ways of determining it*, is the content of a judgment. . . The judgment needs to be expressed by means of a symbol for equality of content, joining the two names together. It is clear from this that different names for the same content are not always just a trivial matter of formulation.

So in that we find a clear definition of the meaning of an identity statement. It is the 'judgment': the same content is given by two names, corresponding to two ways of determining the content. That's what ' $A = B$ ' means for this example.

Evidently, if this is taken as a definition of identity, it is circular, something Russell (1903, 502) spots:¹³

"A is identical with B" means, [Frege] says, that the sign A and the sign B have the same signification. . . a definition which, verbally at least, suffers from circularity.

Wiggins (1965, 51) also spots this problem, arguing that Frege's 2-signs leads to a regress in the following way: If someone asks you for the meaning of $a = b$, the response in line with the 2-signs theory is: a and b have the same content. Wiggins says: 'Unless something is said to justify calling a halt here, the explanation generates a new statement of the same form as the original explicandum'. In other words, the purported meaning of $a = b$ which I just gave as ' a and b have the same content' could be rewritten 'the content of $a =$ the content of b '. We would then need to ask what the meaning of this statement is in the same way as we did about $a = b$. In return, we would regress further: the content of 'the content of a ' = the content of 'the content of b ' and so on ad infinitum.

It's worth noting, as McGinn (2015, 5) points out, that Frege developed the *Begriffsschrift* theory because of the puzzle that he went on to tackle head on in 'On Sense and Reference'. If identity statements relate things designated by names, and ('pure') identity is a relation between an object and itself, then any true identity sentence $a = b$ will not differ from $a = a$: in both cases, '[a] relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but no other thing' (Frege (1892b)). And yet $a = b$ —a

¹³In Frege's defence, it isn't clear that he intends the meaning of an identity statement to be the same as a definition of identity, which Russell possibly thinks he does.

non-tautology—is clearly different to $a = a$ —a tautology. This was why Frege developed the *Begriffsschrift* view of identity as a relation between names. As a name, a is different to b ; $a = b$ is understood as ‘two different names denote the same object’ where $a = a$ couldn’t be given such a meaning as there are not two different names in the first place. The view that identity statements concern a relation between names, then, makes a difference between $a = a$ and $a = b$.

3.3.2 Frege’s ‘On Sense and Reference’ view of identity statements

Frege himself rejected his *Begriffsschrift* view of identity in ‘On Sense and Reference’. Dickie (2008, 269–88) puts it thus: ‘In “On Sense and Reference” Frege rejects this [*Begriffsschrift*] solution on the ground that it treats identity statements as statements about names rather than statements about objects’. Given what we’ve said about Frege’s 2-signs theory, I don’t think it’s a fair reflection of Frege’s own words in the *Begriffsschrift* to say that the problem with the ‘2-signs’ theory is that it is *merely about names*. It’s clearly about names corresponding with the ‘ways of determining a referent’, as the example Frege gave shows (and see section 3.3.1.1).

So what does it mean to say that the problem with the *Begriffsschrift* view is that it is about names, not objects? Frege’s own words criticising his *Begriffsschrift* view in this connection are worth thinking about, even if they are quite hard to grasp:

What we apparently want to state by $a=b$ is that the signs or names ‘a’ and ‘b’ designate the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connexion of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. In that case the sentence $a=b$ would no longer refer to the subject matter but only to its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do. (Frege, 1892b, 56–7)

If identity statements are relations between names, or indeed names corresponding with ‘ways of determining a referent’ (as Frege had it in *Begriffsschrift*, as we saw in section 3.3.1.2), then this is what becomes the ‘subject matter’ of an identity sentence, and this is not ‘proper knowledge’ according to Frege. It is the knowledge that one name co-denotes with another name. But Frege doesn’t think this is what identity statements mean, because he

doesn't think this is sufficiently about objects in the world, and he does think that identity statements are about objects in the world. Someone who hears an identity statement comes to learn something about the world, not just that two names refer to the same thing.¹⁴

In 'On Sense and Reference', Frege introduces the term 'sense', which he says 'serves to illuminate only a single aspect of the thing meant' (Frege, 1892b, 58), 'the thing meant' understood as something mind-independent. 'Sense', therefore, seems not to be a concept or idea, but rather an 'illumination' of a part of something in the world. This, then, allows

¹⁴ This criticism must extend to 'ways of determining a referent' as well, so that Frege can't have been satisfied that identity statements are about names corresponding with ways of determining a referent i.e. Frege presumably didn't want to conclude that $a = b$ can mean that two ways of determining an object correspond to ways of determining one unique object. This seems right; if we employ the technique of substituting identity statements for their meanings/definition, this seems to come out:

- (24) (Ancient Greece c.530BCE)
 Anaximenes: Hey Pythagoras, I got a great view of Hesperus last night. It really is my favourite star. I much prefer it to Phosphorus.
 Pythagoras: I actually just worked out that Hesperus *is* Phosphorus.
 Pythagoras': I actually just worked out that the way of determining the referent for *Hesperus* and the way of determining the referent for *Phosphorus* are two different ways of determining the same object.

Whilst 'Pythagoras'' isn't obviously pragmatically odd, it doesn't seem to be exactly equivalent to 'Pythagoras'. It is about ways of determining a referent, rather than objects in the world. But I confess I'm unclear about the exact difference between 'ways of determining a referent' and the notion of 'sense' that we are just about to encounter. In Frege's *Begriffsschrift* example (section 3.3.1.2), it isn't obvious what the different senses of a point on a line would be, such that points A and B can have different senses. In fact, it is interesting to consider that the example Frege gave in *Begriffsschrift* may admit of an equivalence of meaning between an identity statement and its (*Begriffsschrift*) meaning rather better than (24), especially if we adjust the example slightly. This time, let's say that the way of determining A is also—like B —the consequence of a calculation, just a different one. Let's say that point A is the point of intersection between the circumference of a circle and a straight line running across the diameter of a circle, which is rotated *anticlockwise* 270° around a pivot (P). B is essentially what it originally was, but to be precise here, it is the point of intersection between the circumference of a circle and a straight line running across the diameter of a circle, which is rotated *clockwise* 90° around the same pivot (P) as is used to calculate A . This time, we get the following discourse:

- (25) John: Apparently, there are two points, A and B , which lie on the circumference of this circle. I need to calculate their positions.
 Mary: I actually just discovered that A *is* B .
 Mary': I actually just discovered that the way of determining the referent for A and the way of determining the referent for B are two different ways of determining the same point.

If anything, 'Mary'' is more pragmatically felicitous in the context than 'Mary'. The explanation for this is, I think, that A and B are both points that are worked out on the basis of calculation (the reader may think this is more than a 'slight' adjustment of the original example). In other words, the *Begriffsschrift* view of identity may lend itself better to (25) than to (24) (or classic cases like *Hesperus is Phosphorus*), and indeed it may lend itself better to (25) than the 'On Sense and Reference' view (which certainly seems to lend itself better to classical *Hesperus is Phosphorus* cases, as we're about to see). This is consistent with my general view that identity statements may have a variety of meanings. But, overall, as I said above I'm unclear about this issue, so merely mention it here in case the reader also feels they're missing something.

for an object to have another of its aspects be illuminated by a different sense. To be clear then, senses don't appear to be 'ideas' that we may associate with a word. Frege is quite clear (1892b, 59) that ideas can vary from person to person: 'The idea is subjective: one man's idea is not that of another'. But a sense, '... may be the common property of many people, and so is not a part or a mode of the individual mind. For one can hardly deny that mankind has a common store of thoughts which is transmitted from one generation to another'. So whilst we might have quite different ideas about, say, *Hillary Clinton*, there is a sense—an illumination of an aspect of Hillary Clinton—which is common to all speakers, and which allows successful communication to occur. This follows the interpretation of Frege's notion of 'sense' in McGinn (2015); other authors talk about 'sense' as something apparently mind-internal. For example, Zalta (Zalta, 2017) talks about the 'sense' in terms of 'cognitive significance' and 'ways of conceiving'.

The sense of an expression accounts for its cognitive significance—it is the way by which one conceives of the denotation of the term. The expressions '4' and '8/2' have the same denotation but express different senses, different ways of conceiving the same number. The descriptions 'the morning star' and 'the evening star' denote the same planet, namely Venus, but express different ways of conceiving of Venus and so have different senses.

Ultimately, it may be necessary to resolve this matter if we want to fully understand the nature of identity statements in natural language. For now, I'll follow McGinn's view of senses as relating to mind-external aspects of an object; this seems to make better sense of Frege's dissatisfaction of the *Begriffsschrift* theory of identity statements.

The idea, then, is that two different names will be associated with two different senses in such a way that when they both appear in an identity statement, it is an informative sentence. It tells us that one aspect of an object is an aspect of the same object as another aspect of that object (where the hearer previously thought that the two aspects were aspects of different objects). This appears to overcome the objection that Frege levelled against himself about his own *Begriffsschrift* view of the meaning of identity statements. The objection, recall, was that that view was wrong because it made identity statements *about names* (understood as signs with 'ways of determining' referents). Now, however, identity sentences are about senses of an object, and senses are parts of an object. So now, identity statements really are about mind-external reality, which was what Frege wanted. Some authors (e.g. McGinn (2015)) use the term 'appearance' for 'sense', so that we can say an identity sentence teaches us that two appearances of an object in fact correspond to appearances of the same object. So *Hesperus* was (for the ancient Greeks) the appearance of Venus in the evening, and

Phosphorus the appearance of Venus in the morning.¹⁵

This view of the meaning of identity statements seems quite promising. In the conclusion 3.4, I'll return to show that it seems to substitute effectively for an equative copular.

3.3.3 Giving abbreviations

By way of a slight detour, at the end of section eight of the *Begriffsschrift*, Frege (1879, 12) does acknowledge another way identity statements may be used:

[A] more superficial reason for introducing a symbol for equality of content is that sometimes it is convenient to introduce an abbreviation in place of a lengthy expression; we then have to express equality of content between the abbreviation and the original formula.

So, there is also a definition of an identity statement which is that one is giving an abbreviation, to save us having to repeat a lengthy expression each time.

This is reminiscent, at least, of a notion of identity that Wittgenstein gives in the *Tractatus* (before he gives his own view of identity which I mentioned in section 3.2.3). Bear in mind that Wittgenstein was really addressing issues in formal logic more so than natural language.

4.241 When I use two signs with one and the same meaning, I express this by putting the sign '=' between them.

So ' $a = b$ ' means that the sign ' b ' can be substituted for the sign ' a '.

(If I use an equation to introduce a new sign ' b ', laying down that it shall serve as a substitute for a sign ' a ' that is already known, then, like Russell, I write the equation—definition—in the form ' $a = b$ Def.')

4.242 Expressions of the form ' $a = b$ ' are, therefore, mere representational devices. They state nothing about the meaning of the signs ' a ' and ' b '.

And this certainly seems like a natural meaning. To adapt our example slightly:

(26) (Ancient Greece c.530BCE)

Anaximenes: Wasn't it a clear night last night Pythagoras! I saw the star that's brightest in the evening really clearly.

¹⁵Frege's view of identity in 'On Sense and Reference' is, Russell (1903, 502) claims, similar to his own view of identity in the *Principles of Mathematics*. However, Russell doesn't agree with Frege's view that every proper name has both sense and reference. In his view, only definite descriptions do: '... words such as *John* merely indicate without meaning' ('indicate' is the word Russell uses to mean 'denote'; 'indication' is his translation of Frege's *Bedeutung*, 'reference').

Pythagoras: You mean Hesperus?

Anaximenes: Is that what you call it?

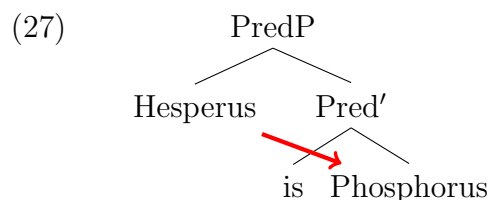
Pythagoras: Yes, we all do Anaximenes, the star that's brightest in the evening is Hesperus.

Anaximenes: Oh, I didn't know that.

So the identity sign here is 'just a permission to substitute one expression for another' (Wiggins, 1965).

3.3.4 Overcoming an expectation of non-identity

The final meaning of identity statements we'll look at comes from Safir (2005). To some extent, understanding Safir's view of identity statements relies on understanding his broader theory of referential dependency and coreference in natural language, which is quite elaborate but I'll introduce it in a very basic way here. Firstly, a noun phrase that *c*-commands another noun phrase cannot referentially depend on that noun phrase (for *c*-command, see footnote 12 in section 2.3.2), something which is the case in *Hesperus is Phosphorus* (let's use a basic Predication Phrase (see 2.3.4) for the sake of exposition):



In (27) *Hesperus* *c*-commands *Phosphorus*. As such, it cannot referentially depend on *Phosphorus*. An example of referential dependence is binding a quantified variable. We can say

(28) [Every boy]_{*i*} thinks Angie loves him_{*i*}

because *him* does not *c*-command *every boy*. But we can't say

(29) *He_{*i*} thinks Angie loves [every boy]_{*i*}

to mean the same thing, because *he* does *c*-command *every boy*, and so cannot referentially depend on it, therefore cannot be quantifier-bound by it. Now, it so happens that when a noun phrase cannot referentially depend on another noun phrase, there is an expectation of non-coreference. We can put this officially (adapted from Safir (2005, 140)):

(30) If X cannot depend on Y, then coreference of X and Y is unexpected.

Coreference is a different kind of relation to dependency; the difference can be seen in the ambiguity of the following example, where *her* can be a referentially dependent or coreferential pronoun:

- (31) Of all the women, only Mara believes Sean loves her
- a. None of the other women believe they are loved by Sean (referentially dependent)
 - b. None of the other women believe that Sean loves Mara (coreference)

In logic:

- (32) a. $\forall x ((x \text{ is a woman} \ \& \ x \neq \text{Mara}) \rightarrow \text{Believe}(x, \text{S loves } x))$
 b. $\forall x ((x \text{ is a woman} \ \& \ x \neq \text{Mara}) \rightarrow \text{Believe}(x, \text{S loves Mara}))$

There are, however, cases where the expectation of non-coreference can be overcome (example originally due to (Higginbotham, 1985, 570)):

- (33) a. A: John is getting up to go.
 b. B: That's not John.
 c. A: Well, he's putting on John's coat.

(33c)'s reply has *he* c-commanding *John*, so *he* cannot depend on *John*. And yet they can corefer. As Safir puts it:

The success of [coreference] in [(33c)] arises because A assumes he knows who John is and that B is a fathead because he does not see that John is who they are looking at. The irony of A's statement in [(33c)] arises from the view that the individual putting on John's coat would have to be John, and that would go *against B's expectation* that John and the guy putting on his coat are not one and the same. (Safir, 2005, 140)

Equatives, then, are another instance where an expectation of non-coreference is overcome. As Safir (2005, 140) says: '[I]n copular sentences like [(34)] the subject *Heracles* c-commands the rest of the sentence, including *Hercules*, and so *Heracles* and *Hercules* should not be allowed to corefer'.

- (34) Heracles is Hercules.

In other words, *Heracles* cannot be referentially dependent on *Hercules* and so the expectation of non-coreference of (30) should kick in. But, as Safir continues (141):

... copular sentences address precisely the expectation of [non-coreference], since they assert that two terms not known to be covalued do indeed have the same value.

He then finishes with the interesting empirical observation: in an example like (35):

(35) Heracles is himself (today).

himself can be dependent on Heracles. The reason for this is that it doesn't c-command Heracles and also that it is the most dependent form (i.e. a reflexive as opposed to a non-reflexive pronoun or fully referential expression, like a proper name) available in the position it is in (this is important to add, as it is the reason why *Hercules* cannot depend on *Heracles* in (34)).¹⁶ However, as Safir observes (141), dependent reference is inappropriate here:

The use of a reflexive is very awkward as an identity statement, because its presuppositions clash—the reflexive established dependent identity whereas the copular asserts independent [coreference]. Such sentences are not ungrammatical, but they are only fully acceptable with idiomatic readings, concerning, for example, Heracles' mental or physical disposition.

So for Safir, an equative is a sentence where the failure of referential dependency (for the post-copula flanker: because it is not a reflexive pronoun and for the pre-copula flanker: because it c-commands the post-copula flanker) induces an expectation of non-coreference. This makes it no different from a sentence like *Heracles likes Hercules*. However, the copular sentence precisely steps in to overcome this expectation of non-coreference. So this is what it means for Safir.

3.3.5 Aspects of identity statements

In the following, I review some ideas which are not explicit statements of the meanings of identity statements, but which are important observations about the nature of identity statements that I think should feature in a statement of their meaning.

3.3.5.1 Genuine reference of the flankers

For Wiggins (1965), distinguishing genuine identity statements from the sort of sentence that is either considered predicational or specificational is about the referentiality of the nominals involved:

¹⁶This latter rule is a part of Safir's system called the 'Form-to-Interpretation Principle' (Safir, 2005, 145)

For a sentence to express an identity ‘is’ or ‘=’ must stand between two noun-phrases which, if they are distinct, are *serving independently of one another to make genuine references*. By this criterion ‘The evening star is the morning star’ and ‘Hesperus is Phosphorus’ do express genuine identities. It is my intention that, *in their normal uses*, such sentences as ‘Darius was the King of the Persians’ (where the question ‘Which individual do you mean to identify by “King of the Persians”?’ is silly) or ‘Plato was the first philosopher to distinguish the copula from the identity sign’ should fail the criterion. (Wiggins, 1965, 42)

Geach, following Wiggins, explicates what it means for an expression to make ‘genuine reference’ (1980a, 1, 60):

[T]he reference of an expression E must be specifiable in some way that does not involve first determining whether the proposition in which E occurs is true. . . it is clearly nonsense to ask which cat “cat” stands for in “Jemima is a cat” . . . I suppose somebody might try saying that in “Jemima is a cat” “cat” stands for Jemima, because the proposition is true. But what the names in a proposition stand for cannot be determined by whether the proposition is true or false: on the contrary, we can determine whether the proposition is true only when we know what it is about, and thus what the names in it do stand for.

As Higgins puts it (Higgins, 1979, 213), ‘[t]he referent is given, so to speak, “in advance of” the assertion about it’.

So there are two interesting insights from Wiggins and Geach’s remarks: (1) We can apply a question test. If it is ‘silly’ or ‘nonsense’ to ask which individual ‘X’ stands for/which individual we mean to identify by ‘X’, then we are not dealing with a true identity sentence. (2) We can’t rely on the truth of a proposition containing ‘X’ to answer this question. We need to be able to answer the question in advance of knowing the truth of the sentence, and in fact, knowing which individual ‘X’ stands for will be necessary to know whether a sentence containing ‘X’ is true or not, in the case of true identity statements. This is not the case for predicationals, as Geach’s example shows. If Jemima is not a cat, then there is no cat such that to ask ‘Which individual does *a cat* stand for?’ can receive an answer. But with an equative, if Hesperus was not Phosphorus then we could have still asked what *Phosphorus* refers to and received an answer. In the case of sentences with attributive noun phrases, we can know

(36) The culprit is a fool.

is true despite not knowing which individual *the culprit* stands for; we merely need to bear witness to the foolish way a crime has been committed, evidenced, perhaps, by the state of the crime scene before us. Again, this is unlike (informative) identities, where we could not know whether the sentence was true or not without knowing what the flankers stood for.

So whilst Wiggins's and Geach's remarks are not an explicit statement of the meaning of an identity statement, they certainly make what I believe are important observations about the 'genuine reference' required of the constituents that enter into identity statements and the way that the truth of these sentences is based on knowing this.

3.3.5.2 Reversibility

In relation to identity statements, Jespersen said:

Where there is perfect identity (coextension) of the two terms connected by *is*, they may change places as subject and predicate; this is what Keats implied in his line: "Beauty is truth; truth, beauty." Jespersen (1924, 153)

Reversibility has subsequently been picked up by the syntax literature and taken to be a criterion for determining whether a copular sentence should be bracketed with the identities or not. Jespersen has a theory of 'specialness' (1924, 75), which seems more or less equivalent to definiteness; a more special word is 'applicable to fewer objects' (the 'most special' words are proper names). In this way, nouns are on the whole more special than adjectives. In 'sentences containing two terms connected by *is*... the member which is most special is the subject... and the less special member the predicative' (Jespersen, 1924, 157). A hidden assumption which I've failed to explicitly locate in Jespersen must be that the more special member precedes the less special member in such sentences. In this way, Jespersen can account for the unacceptability of: **Wise is Socrates*, and can explain the reversibility of identity statements generally; this is where both members are apparently equally special, such that we may well find that it is not possible to say which is the subject and which the predicate, according to Jespersen. Hence either ordering (i.e. either member being the subject or predicate) is allowed.

3.4 Conclusion

If identity is a univocal notion of numerical identity (see section 3.2), it's not obvious that this is something that is ever said about anything else, except in the case where we state what numerical identity is (the fact that every object is self-identical/the fact of the unique

relation which every object has to itself and no other object), or we state the numerical identity of a given object. In the latter case, it isn't obvious that this gives us anything other than a tautology when we say this in the form of a copular sentence (*Hesperus is Hesperus*). Williams (1989, xiii) makes rather the same point:

Identity propositions seem either to state trivially that something is the same as itself, or to state falsely that one thing is the same as another.

Secondly, the ' $\forall P, Px \leftrightarrow Py$ ' side of Leibniz's law, whilst being problematic as a definition of identity, does not in any case seem to be equivalent to the meaning that of an equative copular (see section 3.2.2); substituted in the same contexts, it doesn't obviously perform the same discourse function.

Wittgenstein's attempt in section 3.2.3 to eliminate the identity sign from logic is highly problematic and not at all mainstream. Of course, should identity not be part of the logician's toolkit, then clearly attempts to reduce identity in natural language to logical identity will flounder.

As for the straight meanings of natural language identity sentences that we looked at in section 3.3, these are certainly more promising. Let's put our collection together here:

- Frege's 2-signs theory in the *Begriffsschrift* (section 3.3.1): two names corresponding to two ways of determining a content have the same content.
- Frege 'On Sense and Reference' (section 3.3.2): one aspect (sense) of an object is an aspect of the same object as another aspect of that object (where the hearer previously thought that the two aspects were aspects of different objects).
- Giving abbreviations (section 3.3.3): identity statements 'introduce an abbreviation in place of a lengthy expression' (Frege, *Begriffsschrift*).
- Safir (section 3.3.4): identity statements overcome an expectation of non-coreference.

We also encountered a few criteria that identity statements must meet:

- Wiggins/Geach criteria (section 3.3.5.1): (1) It mustn't be 'silly' to ask which individual/object one of the flanker noun phrases of an identity sentence stands for. (2) The reference of a flanker noun phrase of an identity sentence mustn't be determined on the basis of the truth of the sentence; it must be possible to determine prior to knowing whether the sentence that contains it is true or false.
- Jespersen's reversibility (section 3.3.5.2): identity sentences are reversible.

We already saw in footnote 14 that Frege's 2-signs theory doesn't really give us an equivalent meaning to an identity statement. The 'On Sense and Reference' view seems like it could pass muster though. Let's say, for this example, that the 'sense' of Venus that was associated with *Hesperus* was its being the brightest star in the evening, and the sense of Venus that was associated with *Phosphorus* was its being the brightest star in the morning. In that case it is possible to say:

- (37) (Ancient Greece c.530BCE)
 Anaximenes: Hey Pythagoras, I got a great view of Hesperus last night. It really is my favourite star. I much prefer it to Phosphorus.
 Pythagoras: Actually, Hesperus *is* Phosphorus.
 Pythagoras': Actually, the object which is the brightest star in the evening (and which you call *Hesperus*) is just a different aspect of the object which is the brightest star in the morning (which you call *Phosphorus*)

I'm perhaps twisting the example too much to try and make it work, but it seems to me that 'Pythagoras'' is quite an acceptable paraphrase of the identity statement. Of course, if being called a particular name is a Fregean sense of an object, then we could really just give the paraphrase as:

- (38) Pythagoras'': Actually, the object called *Hesperus* and the object called *Phosphorus* are one and the same/There is only one object which is called both *Hesperus* and *Phosphorus*.

These certainly seem to me, at the very least, to be performing highly relevant functions in the discourse context.

I shall take it for granted without much discussion that the abbreviation function is a possible function for an identity statement, and I gave a version of the equative context in 3.3.3 to show that this seemed plausible.

As for Safir's definition, it performs the right function:

- (39) (Ancient Greece c.530BCE)
 Anaximenes: Hey Pythagoras, I got a great view of Hesperus last night. It really is my favourite star. I much prefer it to Phosphorus.
 Pythagoras: Actually, Hesperus *is* Phosphorus.
 Pythagoras': Actually, Hesperus and Phosphorus are not distinct objects as you seem to think.

But it isn't clear to me that 'Pythagoras'' isn't simply a double negation of 'Pythagoras'

(*Hesperus and Phosphorus are not distinct* \rightarrow *Hesperus is the same as Phosphorus*), which is all well and good, but it hardly explicates an identity statement to say that it is equivalent to such a double negation. It simply begs the question. Unfortunately, in so far as all the equivalents bring in a notion of identity or distinctness, then they are all question-begging to some extent. Perhaps the locution of (38) *There is only one object which is called both Hesperus and Phosphorus* is the best attempt to get round this. But even ‘only one’ is suspect as it only makes sense in the context of more than one, and the notion of plurality, as we saw in section 2.2.3, is standardly understood as presupposing identity.

But my job in this chapter hasn’t been to give a non-question-begging definition of identity statements. The point has really been to discover if there are any well-known accounts of identity statements that would allow us to say that

- (40) a. The culprit is Brian. (specificational)
 b. This is a car. (identificational)

are identity statements in any sense in which philosophers have understood the term. I leave it as an exercise for the reader to go through the list just given and check whether any of these would apply to the sentences in (40). It’s not obvious to me that any of them do. The analysis of identity statements that philosophers have performed seems to be strictly relevant to equative copulars.

In many ways, linguists have followed philosophers in considering specificational sentences as identities. Quine (1943), for example, believed the sentence:

- (41) The number of planets is nine.

was an identity, equivalent to:

- (42) The number of planets = nine

which he proceeded to build an argument about the opaqueness of modal contexts upon: if one then says:

- (43) Nine is necessarily greater than seven.

one may be tempted to deduce, falsely:

- (44) The number of planets is necessarily greater than seven.

Quine’s conclusion is that the modal context introduced by *necessarily* in (43) is referentially opaque so that even substitution of identicals leads to an invalid argument and therefore substitution must be restricted in these contexts. But, as Higgins (1979, 216) points out,

‘Quine and all those who follow him have been systematically misled about [(41)]’s logical form, and it is not an example of what they claim it is. . . [I]t is doubtful whether *the number of planets* has any. . . Referential use at all—it seems rather to be akin to nouns such as *defect* and to have at most the somewhat obscure kind of referentiality associated with indirect questions’.¹⁷ It is interesting to note Higgins’s allusion to a parallel with questions, a view which has only recently been spelled out more formally (e.g. in Heycock (2012)), though which has antecedents in earlier work by Baker (1968). As a simple example of the non-substitutability of *the number of planets* and *nine*, compare:¹⁸

- (45) a. I counted up to nine. vs *I counted up to the number of planets.
 b. Nine is her lucky number. vs *The number of planets is her lucky number.
 c. Nine is the square root of 81. vs *The number of planets is the square root of 81.

Furthermore, where they can appear in the same environment the meaning of the sentence is not the same :

- (46) a. I counted nine.
 b. I counted the number of planets.

(46a) means that one simply counted the series of numbers or counted nine of something, where (46b) has ‘the concealed question type’ equivalent to

- (47) I established how many planets there were by counting them.

Furthermore, there is a curious grammatical difference between (43) and (44) whereby (43) could be rephrased as (48), whereas the equivalent rephrasing of (44) in (49) is ungrammatical:

- (48) Nine is necessarily greater than seven is.
 (49) *The number of planets is necessarily greater than seven is.

Higgins has no explanation for this fact, but simply mentions it as another reason to doubt that the two noun phrases are in any sense identical so that substituting one for the other should be expected to be acceptable, as Quine and many other philosophers have seemed to

¹⁷It may not be fair to attribute this view to Quine himself, as Quine did not think of names or descriptions as referential (i.e. he did not think of the flankers of the copula in (41) as singular terms), so that (41) wouldn’t be an identity statement as such, but would only include an identity between variables (thanks to John Collins, p.c., for pointing this out to me).

¹⁸The asterisks reflects Higgins’s judgements, even though the sentences are not necessarily unacceptable for grammatical reasons.

think.¹⁹

So the upshot of all this is that whatever specificationals are (and this conclusion extends to identificational), they are not identities in any sense that has been laid out by philosophers either studying identity in its ‘pure’ form or the meanings of identity statements. Declerck (1988, 3) seemed to understand this: ‘[Specificationals] are identifying in the sense that they reveal the identity of some entity but not in the sense that they state a relation of identity between two entities. . . [O]nly a couple of linguists. . . have stressed this difference between specificational sentences and identity statements’. And later on (110), he gets it right, I think, when he says that, ‘. . . although identity is logically entailed by [a specificational] sentence, it is not a correct characterization of the specificational act that is performed in such a sentence’. At this point, of course, philosophers and linguists who have claimed these sentences are identities could claim ‘But that’s what we meant! We were interested in a classification of these sentences in terms of whether logical identities were entailed by them or not all along’. But this assumption has never been explicitly stated to my knowledge, and therefore I’ll take it that it is a fair point to say that specificationals (and identificational) cannot be reduced to identities on the basis of any established notions of identity from philosophy. If identity is to have any kind of substantive meaning, it only applies to equative copular sentences.

This concludes the part of this thesis devoted to analysing philosophy to see if there are meanings of predication and identity there which can help us classify copular sentences. Predication, understood as a logical notion, isn’t obviously a meaningful notion of ‘property ascription’ as some linguists seem to think, or indeed a meaningful notion at all. Identity, on the other hand, arguably has been given a meaning by philosophers like Frege, but it only obviously applies to equatives. To try and reduce specificational to identities, as many linguists have argued, therefore, would at least require the linguists stating what notion of identity they are relying on; I can’t find it in philosophy. A popular position (following Russell as we’ll see directly at the start of the next chapter) has been to situate an apparent semantic ambiguity between predication and identity in the copula itself. By now, it should be clear that this ambiguity shouldn’t arise anywhere in the sentence. The next two chapters just make the point that the copula isn’t (even if it was possible, which this chapter and the

¹⁹There are exceptions to this. As Declerck (1988, 110) notes, Linsky (1963) wrote:

Some of the statements that have been counted as identities cannot be interpreted as such. Suppose I explain to my confused son, “Charles de Gaulle is NOT the king of France”. That this statement is not an identity can be shown as follows. From $a = b$, it follows that $b = a$, but from “Charles de Gaulle is NOT the king of France” it does not follow that “The king of France is NOT Charles de Gaulle”. The first of these statements is true while the second is neither true nor false.

last have tried to say it isn't) any kind of a semantically ambiguous linguistic element in any case.

Chapter 4

Is the copula meaningful?

[T]he “is” operator is like adding zero or multiplying by one.

(Geach, 1980b, 182)

4.1 Introduction

The aim of this chapter is to explore the next option we considered in the introduction: can the copula itself be considered a semantically substantive constituent? And furthermore, is it semantically ambiguous in such a way that we can say that it is making the distinctions we seek in order to explain the taxonomy of copular sentences?

One position that has been influential in philosophy of language is that of Russell (1903) and Mill (1856), who believed that *is* in English is ambiguous between a copula, which was essentially a semantically vacuous ‘sign of predication’ (in Mill’s words), and an *is* of identity, which had the semantically substantive meaning of identity (see chapter 3). Russell is often quoted here:

The proposition *Socrates is a man* is no doubt “equivalent” to *Socrates is human*, but it is not the very same proposition. The *is* of *Socrates is human* expresses the relation of subject and predicate; the *is* of *Socrates is a man* expresses identity. It is a disgrace to the human race that it has chosen the same word *is* for those two such entirely different ideas as predication and identity—a disgrace which a symbolic logical language of course remedies. (Russell, 1919, 172)

What is unusual about this remark by Russell is that he should choose *Socrates is a man* as an example of an identity. Post-copular indefinite noun phrases are generally acknowledged

to give rise to predications; as Ludlow says: ‘For linguists it is now standard to think of indefinite descriptions following the copula as always being predicational’ (Ludlow, 2018). About this, Moro says: ‘This reveals that Russell’s idea of identity is really deeply rooted in his logical system and substantially passes over linguistic data’ (Moro, 1997, 254). On multiple occasions Russell also gives as an example of an identity sentence (Russell, 1905, 492):

(1) Scott was the author of *Waverley*.

But (1) is not the sort of identity sentence that Frege was considering in the *Begriffsschrift* or ‘On Sense and Reference’, and linguists fairly universally recognise that it isn’t an equative copular (Ludlow, continuing the quote above: ‘...it is a widespread belief that definite descriptions following the copula are often predicational’ (2018)). Russell seems to speak as though (1) would be as good an example of an identity statement as *Hesperus is Phosphorus*.

In any case, chapters 2 and 3 served the purpose of showing that resorting to logic as a basis for distinguishing copular sentences is not very useful, so this is already an argument against locating a logical distinction in the copula; there isn’t such a logical distinction *anywhere* in a sentence of natural language. In demonstrating that the copula isn’t the sort of thing that could house such an apparently semantically substantive distinction as predication vs. identity (or, at least, the meaning of identity), then we are just providing a further rebuttal of the specific view that locates a logical ambiguity in the copula by saying: the copula isn’t any kind of semantically ambiguous element in any case. So the combination of chapters 2 to 4 is to say that logic does not provide a basis for exhaustively distinguishing the copular taxonomy, and even if it did, it certainly wouldn’t be a consequence of the meaning of the copula.

The issue of whether the copula is ‘semantically substantive’ or not can be broached using methods in linguistics, where there is a well-known distinction between ‘lexical’ and ‘functional’ word classes which is relevant here, and which we’ll look at in section 4.3. Anticipating this, intuitively it seems that some words refer to the world in a way that others don’t. The verb *kiss* seems to refer to the action of kissing, which necessarily involves certain participants with specific roles (a kisser—the agent of the kissing—and a kissee—the patient, receiver, or indeed the ‘experiencer’ of the kissing). For other words, this isn’t so obviously the case. Articles, demonstratives, conjunctions, and modal auxiliaries don’t so obviously refer to the world in the way lexical words do (see section 4.3 below).

This is somewhat reminiscent of the medieval distinction between ‘categorematic’ and ‘syncategorematic’ words. The ‘semantic approach’ to this distinction, notably, makes the point that there are words, ‘which have signification or meaning in isolation from other

words (such as nouns, pronouns, verbs)’ (the categorematic words) and others ‘which have signification only when combined with other words (such as conjunctions, quantifiers, and articles)’ (Uckelman, 2015, 2362). As Henry of Ghent (c.1260) put it:

They are called syncategorematic as if to say ‘consignificant—i.e., significant together with others, namely, with categoremata—not because they signify nothing on their own, but because they have a signification that is not definite but indefinite, a signification whose definiteness they derive from those [words] that are adjoined to them (Kretzmann, 1982, 213).

This is similar to Paul of Pergola’s 15th century definition of those terms which ‘signify *per se*’ and those which don’t (Uckelman, 2015, 2365). But it’s not obvious that the copula receives a categorisation in this approach. The ‘syntactic approach’ to defining these terms specifically refers to subjects and predicates, describing syncategorematic terms as ‘bearing on’ the interrelation between subject and predicate, where the subjects and predicates themselves are categorematic (Uckelman, 2015, 2364). In some ways then, the status of the copula depended on whether one considered it a predicate itself (just the *is* of *Will is tall*) or part of a predicate (the *is tall* part). William of Sherwood’s *Syncategoremata* (c.1250) ranged Latin *est* ‘is’ amongst the syncategorematic words (William of Sherwood, 2012). Interestingly, he draws upon Aristotle as a source (William of Sherwood, 1968), referring to ‘Aristotle’s remark that ‘is’ signifies a certain composition that cannot be understood without the components’; the reference is to *De Interpretatione* 3 (16b23) (Aristotle, 1963): ‘For not even ‘to be’ or ‘not to be’ is a sign of the actual thing (nor if you say simply ‘that which is’); for by itself it is nothing, but it additionally signifies some combination, which cannot be thought of without the components’.

A related question about the copula concerns the extent to which we can think of it as a ‘real’ verb, like *kiss*. Verbs have crucial properties, such as theta role assignment and case assignment, which the copula in English lacks. If the copula isn’t a real lexical verb, then this is further evidence in the case for treating it as a functional word. We’ll address this issue in section 4.4. If the copula isn’t a lexical verb, it isn’t obvious that it would be the sort of thing that would be expected to make the apparently meaningful distinctions that we find in the taxonomy of copular sentences. Whatever else the copula may do, it won’t be responsible alone for the way in which *Hesperus is Phosphorus* is an identity statement, and *Hesperus is bright* is not. It won’t be ambiguous between meaning identity or not, in other words. Meillet was one of the first to recognise that the copula did not have most of the properties of verbs in any particular language (Arche *et al.*, forthcoming)¹

¹Meillet also emphasised the semantic vacuity of the copula: ‘To the extent that a verb figures in a

So what does the copula—the word *be* in English—do? Cross-linguistically, there are a number of different words that approximately correspond to *be* in English: *être* in French, *ser/estar* in Spanish, *essere* in Italian etc. These words perform a variety of functions, and so the general statement of what they do is rather unsatisfying. In a very general way, they are a ‘sign of structure’, though in some languages this ‘structure’ is more obviously semantically substantive than in English. But even in English, the structure may be considered substantive; it includes the idea—which Aristotle subscribed to (den Dikken, 2006a, 267, n.15)—that the copula is a place-holder for tense, which in lexical verbs in English is marked as an affix on the verb itself.² But it goes further than this; it can indicate that other structure is present. Pustet (2003, 2) refers to this function of the copula generally as a ‘syntactic hitching post’ for verbal inflectional categories. In the next chapter, in sections 5.2 and 5.3, we’ll look at two case studies from English and Spanish which will serve to flesh out this conclusion, to say what I mean about the copula being a ‘sign of structure’.

Firstly, however, in the next section I shall briefly address the topic of nominal sentences, a fact about languages which tells against postulating that the copula can be semantically substantive.

4.2 The factual discovery of nominal sentences

Before looking at the distinction between lexical and functional words, a word must be said about what constitutes the most obvious *prima facie* evidence in favour of disregarding the copula as a means of distinguishing the meanings of copular sentences: Moro (1997, 251) puts it thus:

[T]he idea that the copula is not itself a predicate was... factually prompted by the discovery of nominal sentences in Indo-European.

A nominal sentence is simply a copular sentence without a copula. Many languages employ these, and we now know that nominal sentences are widely distributed amongst the language families of the world, and not just Indo-European languages. So for example in Arabic one just says, literally, ‘Mammals taller than crocodiles’:

nominal sentence, it is a grammatical tool entirely denuded of any real meaning of its own’ (Meillet, 1906, 1). Original: Dans la mesure où un verbe figure dans la phrase nominale, il y est un outil grammatical totalement dénué de sens réel qui lui soit propre.

²‘However, the claim that copulas are universally inflectional support elements cannot account for the fact that copulas exist in languages that lack morphological expression of tense, aspect and subject agreement’ (Arche *et al.*, forthcoming).

For example, if I wanted to say that John kissed Mary, I couldn't simply say:

(6) *John Mary.

This seems like a rather obvious point, but it is quite important. Lexical verbs make a substantive semantic contribution. We might say, though, that the copula is making a semantic contribution *of some sort*, and that natural language allows it to be omitted because we have pragmatic inference capacities that allow us to infer the meaning it carries without us needing to actually spell it out explicitly (i.e. to pronounce it) along the lines of the sort of pragmatic inference capacities that have been highlighted in Relevance Theory (Sperber & Wilson, 1995; Wilson & Sperber, 2004). Still, this doesn't apply to lexical verbs. No amount of pragmatic inference is going to allow us to know that (6) could mean that John kissed Mary, rather than that John spoke to Mary. So a difference between the copula and lexical verbs remains.

Some might rejoinder by saying that 'he', the third person masculine singular pronoun in (3) and (5), is also omitted, whilst clearly being semantically substantive. But this is not true. Pronouns are a well-known type of 'reduced referential device', and as such can appear in the most reduced 'zero form' in certain languages (Kibrik, 2011, 74). It is precisely their semantically defective nature (whereby they are dependent on another discourse entity of some kind) that allows them to appear in a zero form. If the copula is like a pronoun in this respect, then this is another argument to consider it semantically defective too. It should also incidentally be pointed out that the copula in many languages is analysed as a pronoun, not a verb (Pustet, 2003, 45).

4.2.1 More details about copula variation cross-linguistically

The precise distribution of the presence vs. absence of the copula is subject to cross-linguistic variation. Here below, I'll just give a few details of the kind of variation different languages exhibit. The point of doing this is merely illustrative. Two things should be apparent: (1) Languages often don't use copulas where English does, (2) The conditions under which different languages do or do not use a copula are not uniform cross-linguistically, and certainly don't in any obvious sense correspond to the the distinctions of the copular taxonomy. A view that follows from describing the copular taxonomy in terms of the logical notions of predication and identity is that the *be* in predications is a semantically vacuous copula whereas *be* in identity sentences is a semantically substantive verb, meaning identity. As such, a natural conclusion is that *be* can be omitted when it is a copula, but not when it is the *be* of identity. Looking at natural language variation gives little support to this

conclusion, and hence to the view that copular sentences are ultimately well-described in logical terms. In 4.2.3, I discuss the apparently most direct evidence for the view that the copula has a substantive identity meaning.

When discussing the cross-linguistic variation below, I'm allowing myself a certain amount of leeway in my use of the term 'copula'. Strictly speaking, whatever definition of the copula in English we come up with would not necessarily apply to any linguistic item in any other language. It would therefore be more accurate to talk of other languages' *be*-like constituents, rather than calling them uniformly 'copulas'. This is true about languages generally. English *a* and French *un* are both called 'indefinite articles', but *un* spells out a Number Phrase in French, not in English (Roy, 2013), which is why it is homophonous with the number one in French, but not in English. Similarly for the 'definite article'; *the* and *le* can have quite different uses in English and French, for example, hence the difficulty language learners encounter when they realise their 'definite article' cannot be used before plural count nouns to refer generically (one cannot say *the dogs* to refer to dogs in general in English—unlike French where *les chiens* can refer to dogs in general and *chiens* is simply considered ungrammatical in most contexts). And so on through much of the rest of a language's functional and lexical words.

With that caveat in mind, on to some details of cross-linguistic variation: A common property of languages is to omit the copula in the 3rd person only, such as Hungarian:

- (7) a. Ö diák. Hungarian
 lit. He a student
 'He is a student.'
 b. Én diák vagyok.
 lit. I a student am
 'I am a student.'
 (examples from Ferguson (1971, 142), no glosses)

In Russian a copula is also not used in the first person:

- (8) Ja student. Russian
 lit. I student
 'I am a student.' (example from Ferguson (1971, 142), no gloss)

Some languages limit the use of copulas to adjectival predicates, but not nominal ones. So in Haitian Creole the copula *sé* can be absent with an adjective *lazy* in (9a), but must be present with nominal 'stallions' in (9b):

- (9) a. Chwal yo parésé. Haitian Creole
 ‘The horses are lazy.’
 b. Chwal yo sé étalô.
 ‘The horses are stallions.’ (examples from Ferguson (1971, 142), no gloss)

Some languages, like Bengali, use a copula in dependent clauses (10a), but not in main clauses (10b):

- (10) a. Se jodi chatro hoě... Bengali
 he if student copula
 ‘If he is a student...’
 b. Se chatro.
 ‘He is a student.’ (examples from Ferguson (1971, 142), my gloss)

Perhaps most notably, when the tense of the sentence is anything other than the present tense, copulas are generally used in languages that otherwise omit them in the present tense. So in Arabic, we see copulas in the past tense (copula in bold):

- (11) ‘Axī **kāna** tilm.ī an. Arabic
 ‘My brother was a student.’ (example from Ferguson (1971, 142), no gloss)

This is by way of comparison with the present tense Arabic sentence in (2) where there is no copula. In fact, another example allows us to make the stronger claim that past tense in Arabic in fact *requires* a copula:

- (12) *(kāna) r-rajul-u fī d-dār-i. Arabic
 was the-man-Nom in the-house-Gen
 ‘The man was in the house.’
 (example from Bahloul (2006, 508))

This is also seen in African American Vernacular English (AAVE):

- (13) a. She crazy. AAVE
 ‘She is/*was crazy.’
 b. She *(was) crazy.
 ‘She was crazy.’
 (example from den Dikken and O’Neill (2017, 15))

Another example from Arche et al. (forthcoming) shows that copulas are sensitive to polarity, as well as tense, in Arabic:

- (14) a. L-madi:natu ?a:miratun. Arabic
 the-town busy
 ‘The town is busy.’
 b. Laysat l-madinatu ?a:miratun.
 is.not the-town busy
 ‘The town is not busy.’

As the above shows, the literature is often not completely explicit about when nominal sentences are merely possible or in fact necessary. This is important to know. For example, is there a possible sentence in Arabic which is a version of (2) except with a copula present? If so, it would then be important to check whether this is a case of true optionality or whether the presence of the copula is making some kind of semantic distinction. As I say, the literature is often not explicit about this. But these would be important facts to know in the context of the debate about ‘true optionality’ in general (Biberauer & Richards, 2006): the issue of whether natural language ever contains minimal pairs that are semantically entirely equivalent. Later in this chapter, in section 5.2, we’ll look at the case of English minimal pairs like:

- (15) a. I consider Mary intelligent.
 b. I consider Mary to be intelligent.

Whilst they may appear to be semantically equivalent, there is in fact a subtle difference that shows up in the presence of quantification. Benveniste (1966) also explicitly considers the issue of whether copular and nominal sentences are semantically the same or different, and we’ll turn to his thoughts about this matter now.

4.2.2 Semantic differences between nominal sentences and copular sentences according to Benveniste

As mentioned above, Benveniste considers that there are systematic semantic differences between copular and nominal sentences in languages which appear *prima facie* to allow free variation between them. Benveniste based his study on ancient languages, attempting to discern semantic differences from the uses of nominal and copular sentences in classical texts. Latin, for example, allows for minimal pairs like:

- (16) a. Omnia praeclara rara. Latin
 all excellent rare
 ‘All excellent things are rare.’

- b. Omnia praeclara rara sunt.
 all excellent rare are
 ‘All excellent things are rare.’

Is there any real difference between (16a) and (16b)? (Benveniste, 1966, 159) puts it thus:

In the nominal sentence, the assertive element, being nominal, *can not* carry the specifications that a verb can: modalities of tense, person etc. The assertion has a timeless, impersonal, non-modal character—in short, it expresses the assertive element’s intrinsic semantic content alone. A second consequence is that this nominal assertion can not either participate in the essential quality of a verbal assertion, which is to place the time of an event in relation to the time of the utterance about the event by a speaker. The nominal sentence in Indo-European asserts a certain ‘quality’ (in the most general sense) as belonging to the subject of the utterance, but with no temporal or other kind of determination or any kind of relation to the speaker.³

But this doesn’t mean that a nominal sentence is ‘incomplete’ in any sense. Benveniste is explicit: ‘It is as complete as any verbal utterance’⁴ (159).

Crucially, Benveniste was explicit about the way that we shouldn’t consider that e.g. Latin

- (17) Omnis homo mortalis. Latin
 Every man mortal
 ‘All men are mortal.’

is the zero-copula form of *Omnis homo mortalis est*. What is his evidence for this? Well, he points out that ‘*esti*’ (which I take to be a proto-Indo-European form that was the ancestor of the Latin third person singular present copula *est* and its Romance language derivatives) was a ‘real’ lexical verb with a ‘definite lexical meaning, before it fell—over the course of a long historical development—to the rank of a ‘copula’⁵ (Benveniste, 1966, 160). Benveniste

³Original: Dans la phrase nominale, l’élément assertif, étant nominal, n’est pas susceptible des déterminations que la forme verbale porte: modalités temporelles, personnelles, etc. L’assertion aura ce caractère propre d’être intemporelle, impersonnelle, non modale, bref de porter sur un terme réduit à son seul contenu sémantique. Une seconde conséquence est que cette assertion nominale ne peut pas non plus participer à la propriété essentielle d’une assertion verbale, que est de mettre le temps de l’événement en rapport avec le temps du discours sur l’événement. La phrase nominale en indo-européen affirme une certaine ‘qualité’ (au sens plus général) comme propre au sujet de l’énoncé, mais hors de toute détermination temporelle ou autre et hors de toute relation avec le locuteur.

⁴Original: Elle est aussi complète que n’importe quel énoncé verbal.

⁵Original: ...sens lexical défini, avant de tomber—au terme d’un long développement historique—au rang de ‘copule’.

speculates that it had a meaning closer to ‘grow’: ‘It is no longer possible to detect this meaning directly, but the fact that **bhū-*, ‘to put forth, to grow’, has supplied some of the forms of **es-* gives us a glimpse of this meaning’⁶ (Benveniste, 1966, 160). I think Benveniste is slightly confused in this last quote. Both **bhū-* and **es-* are Proto-Indo-European forms that have fed into the copula systems of Indo-European languages. **bhū-* (Ringe, 2006, 10), for example, is considered to be the ultimate root of English *be*, *been*, whereas **es-* is the Proto-Indo-European form which is the root of Latin *est* (and its Romance language derivatives) (Ringe, 2006, 35). So I think Benveniste really means that **bhū-* provided some of the forms of the copula in different Indo-European languages (i.e. the infinitive and past participle forms in English).

Basing himself on ancient Greek, as noted, Benveniste lists every nominal sentence in Pindar’s *The Pythian Odes*, examples like:

- (18) τὸ δε μόρσιμον ου παρφυχτόν. ancient Greek
 ‘Destiny remains inevitable.’ (no gloss given by Benveniste)

This leads him to further conclusions that elaborate his semantic distinction between nominal and copular sentences. For example, nominal sentences ‘always serve to make assertions with a general, sententious, character’⁷, and he notes that the French linguist Meillet had already observed that nominal sentences express ‘general truths’. The ‘narration of a fact’ or the description of ‘a fact in its particular circumstances’⁸ can, by contrast, only be done by a verbal phrase (with, in Greek, *εστι*). Benveniste (1966, 165) says that nominal sentences and sentences with Greek *εστι* aren’t part of the same ‘register’, ‘[t]he first relates to speeches, the second to narratives. The first poses absolutes, the second describes a situation’⁹. Nominal sentences are entered into discourse ‘to influence and convince, not to inform’¹⁰. Benveniste contrasts Pindar’s odes with Herodotus’ *History*, deliberately chosen as an example of narrative prose, and duly finds far fewer nominal sentences (fewer than ten examples, all of which express ‘general truths’) and an ‘abundant’ use of *εστι*, aligning with his description of the semantic differences between nominal and copular sentences.

⁶Original: Il n’est plus possible d’atteindre directement ce sens, mais le fait que **bhū-*, ‘pousser, croître’, a fourni une partie des formes de **es-* permet de l’entrevoir.

⁷Original: ...sert toujours à des assertions de caractère général, voire sentencieux [where I believe ‘sententious’ is intended in the sense of ‘aphoristic’ rather than ‘pompous’].

⁸Original: un fait dans sa particularité.

⁹Original: La première est du discours; la seconde, de la narration. L’une pose un absolu; l’autre décrit une situation.

¹⁰Original: ... pour agir et convaincre, non pour informer.

4.2.3 The presence/absence of the copula in predicational vs. equative copular sentences

So much for the fact of nominal sentences and the apparent semantic differences between nominal and copular sentences in languages which have both (like ancient Greek). In fact, it is quite a standard view in philosophy and linguistics that the copula of predicational copular sentences is basically semantically vacuous (even though predication itself is claimed to be an apparently semantically substantive ‘property ascription’ relation). This contrasts with the *be* in equatives which is held to be semantically substantive. This was the view we saw expressed by Frege in 2.2.7. A piece of data that is routinely brought out in support of that view is the following from English (examples from den Dikken and O’Neill (2017)):

- (19) Predicational
- a. I consider Dr. Jekyll (to be) handsome.
 - b. Dr. Jekyll seems (to be) handsome.
 - c. With Dr. Jekyll (being) so handsome, we will easily win the contest.
- (20) Equative
- a. I consider Dr. Jekyll *(to be) Mr. Hyde.
 - b. Dr Jekyll seems *(to be) Mr. Hyde.
 - c. With Dr. Jekyll *(being) Mr. Hyde, our mystery is solved.

and from Hebrew (Rothstein, 2001, 207):

- (21) a. Dani (hu) more.—predicational Hebrew
 Dani COP.M.SG teacher
 ‘Dani is a teacher.’
- b. Dani *(hu) mar Yosef.—equative
 dani COP.M.SG Mr. Yosef
 ‘Dani is Mr. Yosef.’

The point that is then made is that the copula must be obligatorily present in equatives, thereby indicating that it is semantically substantive (and therefore not omissible) unlike the copula in predicationals which is optional.

Frankly, I don’t know exactly how to explain the data above and it does matter for me; it wouldn’t be consistent with this thesis to allow for a language to syntactically mark the predicational vs. equative distinction. One reason is that I conclude later in the thesis that *all* copular sentences are syntactic predications. To be predicational may, I argue later in chapter 8, relate to an information structural fact (interacting with the way the flankers in a

copular sentence refer). Conceivably this is somehow unmarked, or allows for the optionality of the copula, where equatives—which generally put nuclear tone on the copula itself—are different and cannot be so unmarked. But this is completely speculative. The wider point of this section 4.2 has been to make the basic point that copulas are required for all kinds of reasons. To try and capture the variation of the examples we’ve seen in this section in terms of the copula—when it is obligatorily present—‘meaning identity’, doesn’t seem like a very promising avenue. In the examples in 4.2.1, it is clearly because of the person or tense of the sentence that a copula is obligatory, not the identity meaning of the sentence. Finally, the examples in (42), from Russian and Warlpiri, show that using a copula in equatives is not in fact cross-linguistically universal (as den Dikken and O’Neill (2017) seem to think). The examples in (42) are equative nominal sentences. However, despite these facts, it would still remain for me to properly explain the data in (21) and the English data preceding it. This would really be first-order linguistic work, certainly requiring a greater amount of data from Hebrew in the case of (21). So I shan’t attempt that here. I’ll merely predict that it should be possible to explain these facts in ways that do not require resorting to predication vs. identity as primitives.

4.3 Lexical vs. functional categories

In this section, I’ll briefly elaborate on the difference between lexical elements and functional elements in natural language. From this, we should be able to discern the status of the copula. Haegeman and Guéron (1999, 20–1) say:

We will also make a distinction between lexical elements and functional elements. The former are items such as nouns (*John, woman*), verbs (*sleep, eat*), adjectives (*big, happy*), adverbs (*quickly*) and prepositions (*in, under*). Lexical elements contribute to the descriptive content of the sentence. Functional elements are elements such as modal auxiliaries (*will, shall*), articles (*the, a*) and conjunctions (*that, if, for*). They do not contribute to the descriptive content of the sentence.

Later (59–61) Haegeman and Guéron (1999) also include demonstratives and possessives alongside articles, which they bracket under a broader ‘Determiner’ (‘Det’) heading.

In terms of articulating the difference between lexical and functional items, they say: ‘First, it is possible to list all the elements that belong to Det, while it would be well nigh impossible to list all English verbs. One can create new verbs quite easily when they are needed to express new notions. But the class of determiners is fixed’ (Haegeman & Guéron, 1999, 61). This then overlaps with the notions ‘open class’ and ‘closed class’ which specifically

refers to this aspect of word classes to be indefinitely expandible, as the lexical categories are, or fixed in terms of their members, like the functional categories. Prepositions are a ‘marginal’ category in this respect. Haegeman and Guéron say: ‘It is true that there is a relatively restricted number of prepositions and that new prepositions are not created every day’ (61), before giving an example: ‘[S]uppose that it becomes important to be able to refer to a relation in which an element *X* is three centimetres to the right of an element *Y*. One might invent the preposition *agrip* for this, and it would not be unnatural to say *The book is agrip the box*’ (61).

Haegeman and Guéron (1999, 61) rely on their readers’ intuitions about ‘lexical content’:

There is an important semantic contrast between words from the categories N, V, A, Adv, and P, and words belonging to the class of Determiners. Elements of the categories N, V, A, Adv, and P have a lot of lexical content. The difference between (108a) and (108b) is due to the choice of verb and the events described in these sentences are quite distinct:

- (108) a. The woman **met** the president.
 b. The woman **hit** the president.

On the other hand, the choice of *the* versus *a* in (108) does not give rise to a change in the state of affairs depicted by the sentence... Similarly, replacing an article by a demonstrative determiner does not alter the descriptive content of the sentence.

They then give some examples in order to show that replacing nouns, prepositions, adjectives, and adverbs has an impact on the ‘type of event depicted by the sentence’:

- (22) a. The **woman/man/sailor** met the **president/doctor/butcher**.
 b. I met Georgina **on/under/behind** the railway bridge.
 c. Louise is a **brave/ambitious/generous/sexy** woman.
 d. Thelma corrected the text **painstakingly/carefully/quickly/generously**.

This all leads to the conclusion that lexical categories have ‘lexical content’ and functional categories have ‘functional content’. The functional categories are not semantically completely denuded: the difference between *the woman* and *a woman* is to do with the way that *the woman* (in, say, *The woman hit the president*) refers to a referent that is salient in some way, rather than *a woman* (in, say, *A woman hit the president*) which would be newly introduced (von Heusinger, 2013). With demonstratives, *that* and *this* contrast with respect to how near or far a referent is to the speaker in some way (Langacker, 1987, 91–2).

Lexical content for Haegeman and Guéron (1999) involves: ‘descriptions of distinct events’, ‘changes in a state of affairs’, ‘descriptive content’, ‘types of event depicted by the sentence’. I shan’t really elaborate on these grounds for the notion of lexical content; descriptions of events and states of affairs seem to be at the heart of lexical content in a way that they aren’t for the functional categories.¹¹

Another way in which functional categories differ from lexical categories relates to their phonetics. Functional categories can often be ‘phonetically reduced’ (Haegeman & Guéron, 1999, 62). They give the example of the definite article in French, showing that it often forms a phonetic word with the noun it is associated with to its right or with a preposition to its left:

- (23) a. **le** + armoire → **l’**armoire
 b. **à** + **le** garçon → **au** garçon

From the above, we have a number of criteria which may help us determine where the copula fits in this functional vs. lexical classification, and here I shall briefly present the case for considering the copula as a functional element; in particular, it appears close to a kind of auxiliary verb, one of the standardly accepted functional categories.

4.3.1 Copulas are inflection-like in many languages

Firstly, whilst auxiliaries in English seem superficially like verbs in that they are separate words that inflect for tense e.g. *will* vs. *would* in (24a), this is not cross-linguistically universal, and in many other languages, auxiliaries are simply expressed with verbal inflection, as in French, (24b):

- (24) a. Thelma **will/would** meet the president.
 b. Thelma **recontrera/ait** le président.

This is unlike lexical categories, that do not appear as inflectional affixes in any language (they are the stems, the things that get inflected). What about the copula? Recall the caveat above (section 4.2.1) about the way that ‘copula’ is not a cross-linguistically uniform category. The best we can say about other languages is that they have elements that seem to perform the function of the *be* copula in English and we often don’t yet know how else to describe them. Some languages exhibit affixes that so far can’t be described as anything except ‘copula affixes’. Affixes are not the same thing as inflections exactly, but they can be suggestive of

¹¹Rather as with property ascription (see section 7.3), I suspect these may be metaphysical notions ultimately.

the way that the copula is more like an inflection in some languages than a lexical category. Turkish, for example, realises certain inflections—such as person inflections—directly on the ‘predicate nucleus’ (Pustet, 2003, 4). So this is rather like a sentence of English where the third person singular *-s* inflection is added directly to *happy* in *Bill is happy* rather than appearing on the copula, as it in fact does. Turkish has a ‘copula affix’ *-y*:

- (25) (Ben) satıcı-y-ım. Turkish
 (I) seller-COP-1sg.
 (example from Pustet (2003, 4))

The fact that the copula affix attaches directly to the predicate along with other inflectional affixes (person and number) is suggestive of the way that we can analyse the copula as, at least, inflection-like in Turkish, something that makes it seem more like a functional than a lexical category.

4.3.2 Copulas are, or are part of, a closed class

Above we noted that lexical categories are open class where functional categories are closed class. If we consider ‘copula’ to be a category in itself (like ‘auxiliary’), we can ask about this category whether it is open or closed class. Here the traditional grammars are quite clear. In English, the most common copula by far is *be*, however a restricted set of other words is often included. According to Quirk and Greenbaum (1973, 352–3), they fall into two main categories according to whether the predicate is that of ‘current attribute’ or ‘attribute resulting from the event described in the [copula] verb’. They list the following, saying ‘[m]ost of them are used only with a subject complement that is an adjective phrase or a noun phrase with gradable head noun’.

- (26) a. ‘Current’ copulas: *appear, feel, look, remain, seem, smell, sound, taste*
 b. ‘Resulting’ copulas: *become, get* (chiefly informal), *go, grow, turn, make*

To be clear, these words above can appear in copular sentences, i.e. where they join a subject and predicate directly:

- (27) a. Mary appeared happy, John felt sad, Bill looked a fool, Sara remained on strike,
 Tom seemed annoyed, The melons smelt delicious etc.
 b. The work became a burden, Jamie got drunk, The fans went crazy etc.

Clearly the vast majority of lexical verbs cannot be used in this way:

- (28) a. *Mary kissed happy, *John spoke to sad, *Bill gave a fool, *Sara took on strike,

*Tom bought annoyed etc.

The examples above should not be confused with ‘depictives’—*Malcolm left home young, Alan walked around naked*—or ‘resultatives’: *Malcolm hammered the metal flat* (Jackendoff, 1990). Depictives and resultatives are a kind of secondary predication. By secondary predication, we can understand that a property is predicated of a participant *in an event*; the predication of (27) does not involve event participants in an obvious way (such sentences are usually thought of as describing ‘states’; admittedly, some authors such as Roy (2013) following Parsons (1990) consider states a type of event). Plausibly, some of the examples in (28) can be construed as depictives/resultatives, if we allow a marginal intransitive use of e.g. *kiss* and *bought*. It’s difficult to avoid such construals in so far as lexical verbs often describe actions/events, more so than the kind of stative verbs that are categorised as copulas.

So copulas appear to be a closed class. This lends further support to viewing them as a functional category.

Whilst it is not standard to consider the copula as a type of auxiliary verb in English, there is of course a suggestive homophony between the copula and the progressive and passive auxiliary. Auxiliaries are also a closed class: the number of auxiliaries in English is generally restricted to the modal auxiliaries (*will, would, shall, should, can, could, may, might, must, have to* and a few others like *used to* are sometimes also included) plus the perfective (*have* in e.g. *has/had met*), progressive (*be* in e.g. *is meeting*), and passive (*be* in *was eaten*) auxiliaries, and *do* in questions and negation (*Did you take it? No I didn’t touch it!*) (Quirk & Greenbaum, 1973).

4.3.3 Copulas are semantically functional-like

In a purely semantic way, Haegeman and Guéron (1999, 62) say about auxiliaries that ‘[they] do not depict situations; rather, they express certain modifications of time, mood or aspect, with respect to the type of situation expressed by the verb and its arguments’ by way of justifying the status of auxiliaries as functional elements. The same could be said of the copula in English with the adjustment that it expresses modifications of time, mood, and aspect with respect to the predicate, whatever it is, be it an adjective, nominal, or prepositional phrase. But the copula itself does not ‘depict a situation’; it isn’t itself the predicate that is applied to a subject, it merely allows certain modifications of the predicate in terms of time, mood, and aspect.

4.3.4 Copulas are distributed like auxiliaries

Another way in which copulas are similar to auxiliary verbs, a classic functional category, is their distribution with respect to adverbs, negation, and question words. Haegeman and Guéron (1999, 84ff.) note the following pattern: auxiliaries—unlike lexical verbs—precede a sentence-medial adverb:¹²

- (29) a. Thelma will often read this book.—modal auxiliary *will* precedes adverb
 b. Thelma has already read the book.—perfective auxiliary *has* precedes adverb
 c. *Louise reads often novels.—lexical word cannot precede adverb (cf. *Louise often reads novels*)

What about the copula?

- (30) Louise is already bored.

Copulas appear to pattern like auxiliaries. **/?Louise already is bored* is at the very least a marginal way of saying (30). Negation also provides a distributional test. Auxiliaries precede *not*:

- (31) a. Thelma will not write any books.
 b. Thelma has not written any books.

This is not possible for lexical verbs:

- (32) *Thelma writes not any books.

The copula, again, patterns with auxiliaries:

- (33) Thelma is not bored.

Finally, in question formation auxiliaries precede subjects:

- (34) a. Will Thelma write any books?
 b. Has Thelma written any books?

This is not possible with lexical verbs:

- (35) Writes Thelma any books?

Again, the copula patterns with auxiliaries:

¹²Sentence-medial adverbs are to be distinguished from sentence-initial or sentence-final adverbs, to which the following constraints do not apply.

(36) Is Louise bored?

4.3.5 Copulas phonetically reduce like functional words, except in equatives

Finally, recall the phonetically reduced nature of functional elements we noted above in 4.3. This applies to auxiliaries generally, which almost always in speech and informal writing are adjoined to subjects (37a); this applies equally to the copula *be* (37b):

- (37) a. She'll meet the president/She's met the president/She's meeting the president.
 b. Mary's a pilot/John's drunk/Bill's in the garden/That's a car/The culprit's Brian.

Equatives don't seem to allow such contractions so felicitously:

- (38) a. ?The evening star's the morning star.
 b. ?Holland's the Netherlands.

This is presumably because in such sentences the obvious context for them is such that there would be nuclear tone on the copula itself, and prosodically prominent linguistic elements cannot be phonetically reduced:

- (39) A: You don't seem ready at all.
 B: I **am** ready!
 B': ?**I'm** ready! ?I'm **ready**!

So, to be clear, I don't take the impossibility of reducing the copula in an equative as evidence that there it is a lexical verb. Rather, I believe it remains the same copula but that equatives (unlike the other copular sentences) generally require some kind of contrastive accent on the copula itself.

4.4 Is the verbal copula even a real verb?

The previous section has anticipated this section to some extent. In the previous section, we tried to show that the copula was a functional (as opposed to lexical) word, and specifically how it was similar to an auxiliary verb, a well-accepted functional category. As such, it wouldn't be expected to play the apparently lexical role of distinguishing the meaning of predication from identity. The point of this section is to say that copulas are not just a functional category—unlike lexical verbs—they aren't even really proper verbs at all.

4.4.1 Does the copula assign theta roles?

For many, having a theta grid is a criterion of being a verb, moreso than merely being a dummy element that hosts certain inflectional elements, like tense, mood, and aspect. Given that since Higginbotham (1985), it has been common to consider that all lexical categories (verbs, nouns, adjectives, and prepositions) have theta-grids, some go further e.g. Pereltsvaig (2001, 50) says, ‘the characteristic property of lexical categories which distinguishes them from functional categories is their association with theta-grids’. The way in which nouns can be said to assign theta roles usually arises from comparisons like the following (originally noticed by Chomsky (1970), though not in the context of theta theory):

- (40) a. John criticized the book.
 b. John’s criticizing the book

where the nominal form in (40b) inherits the same argument structure as the verbal form in (40a). Once we add theta roles to argument structure, we can say *John* and *the book* are held to be thematically related to the noun *criticizing*.¹³ A ‘criticizing’ seems to require a criticizer and a thing criticized in the same way.

Does the copula assign theta roles? Haegeman and Guéron (1999, 39–40) give the following examples:

- (41) a. I consider [that his reply is inadequate].
 b. I consider [his reply inadequate].

They note that in (41a) *consider* takes two arguments *I* and *that his reply is inadequate*. The question arises for *his reply*: is this an argument, and if so what is it the argument of? The options are the copula *is* or the adjective *inadequate*. But the same question arises for (41b), where there is no copula, ‘but semantically the interpretation of [(41b)] is similar to that of [(41a)]. . . We are thus led to conclude that adjectives can also assign thematic roles’ (Haegeman & Guéron, 1999, 40). (In the next chapter, in section 5.2, I’ll show that there are conditions in which sentences like (41) can be differentiated with respect to the scope relations they permit.). So *inadequate* takes *his reply* as an argument and assigns a theta role to it in both sentences in (41), in line with what we just said about all lexical categories being theta role assigners. So crucially, the examples in (41) show that, in so far as *his reply* receives a theta role, it cannot be assigned by the copula; the copula is absent in (41b) but the sentence is still acceptable because *his reply*’s theta role is assigned by the adjective.

Examples like (41) show a copula in a typically predicational context. What about the

¹³The form *criticizing* in (40b) is a gerundive, the categorical status of which has in fact been debated (Hudson, 2003).

so-called *be* of identity in equatives? Doron (1983) and Rapoport (1985) claimed that the difference between *be* in predicational sentences and *be* in equatives relates to its role as a theta-role assigner only in the latter case. But Rapoport (1987) retreats from this claim. Generally, as Rapoport observes, ‘it is not exactly clear which theta-roles the copula would assign, and how the identity relation would be derived from that assignment’ (141, see also Heggie (1988)).¹⁴ Then her argument is essentially the same as Haegeman and Guéron’s just above. She says that if *be* is required to assign theta-roles, then this should be a cross-linguistic universal (‘given that semantics is universal’ (Rapoport, 1987, 140)), and yet plenty of languages have equatives in the absence of a copular, e.g.

- (42) a. Ivan eto tot samyj čelovek Russian
 Ivan this-n this-m very man
 ‘Ivan is this very man.’
 b. paddy yamingi. Warlpiri
 ‘Paddy is Yamingi.’

Finally, Rapoport makes the following point: let’s assume that *be* assigns two theta roles in equatives and none in predications (it can’t assign theta roles to predications because the predicative flanker already assigns one to its subject according to Rapoport (and Haegeman and Guéron, just above); therefore, if the copula also assigned one to the subject, the subject would carry two theta roles, in violation of the theta-criterion, see section 2.3.1). This, however, is ‘...rather an odd view of optionality of theta-role assignment’ (141–2). It is unusual, Rapoport feels that either both theta roles or neither must be assigned. Normal optionality should allow for a theta role to be ‘independently optionally assigned’.

An argument that there are two *bes*, one of which assigns theta roles is given in Safir (1985, 117–8). Safir takes the contrast between:

- (43) a. I thought John to be a fool.
 b. I thought John a fool.

and

- (44) a. I thought Shakespeare to be the author of *The Tempest*.
 b. *I thought Shakespeare the author of *The Tempest*.

¹⁴Rapoport seems to assume that theta role assignment should do the job of equating the nominals of an equative; why she makes this assumption is not entirely clear to me. Perhaps if we compare the copula in an equative with a verb like *kiss*, we can see more clearly. The theta roles that *kiss* assigns are, essentially, the ‘kisser’ and the ‘kissee’. But it still seems that we need the verb itself to give us the meaning of the action of kissing, in the same way as it could be argued that we need the copula to give us the meaning of equation, beyond any theta roles it may assign.

to be a result of the theta-criterion being violated. This is because he believed that (44) contains an ‘identificational *be*’ which assigns two theta roles to the pre- and post-copula arguments. As such it needs to be present in (44). In contrast, Safir considered that (43) contained a ‘predicational *be*’ which is semantically vacuous, and which selects a small clause (see section 4.4.2 just below) where the predicate (*a fool* in (43)) is a theta role assigner which assigns its role directly (to *John* in (43)) without the copula interceding in any way (as we’ve already stated is possible on the part of all lexical categories in section 4.4.1).

Safir says that, ‘it is generally the case that definite NPs are not used predicatively’ (Safir, 1985, 118). But this is in stark contrast to the quote earlier in section 4.1 from Ludlow that ‘it is a widespread belief that definite descriptions following the copula are often predicational’ (Ludlow, 2018). Mikkelsen (2011, 1809) says: ‘The predicate complement [of a predicational copular sentence] can be an AP... or a PP, DP or NP’—the DP examples she gives are:¹⁵

- (45) a. Sylvie is an architect.
 b. Sylvie is the architect on that project.
 c. Sylvie is my friend.

On balance, I believe the syntax literature generally accords with Ludlow and Mikkelsen in allowing that DPs are able to serve as copular predicates, in which case they too should be able to assign a theta role directly to the subjects in (45). So whatever the reason for the apparent unacceptability of (44b), I shall not view it as a theta criterion violation here.

The evidence in (44b) has been regularly taken to suggest that there is a distinct ‘equative copula’ which is required to express a semantically substantive identity meaning (as opposed to the semantically vacuous predicational meaning).¹⁶ But as Heycock and Kroch (1999, 381) pointed out, there are examples in English of copula-less equative clauses, when they are embedded under *make*:

- (46) a. But if what you say is true, that would make the real murderer John!
 b. But if what you say is true, that makes your attitude toward Jones my attitude toward Davies!

Furthermore, arguing the other way, equative semantics can also be found in the presence of verbs other than the copula. Heycock (1994) shows that equative small clauses occur with *remain* and *become*:

¹⁵Mikkelsen is interpreting DP as any noun phrase with an article, as per the definition I gave in chapter 2, footnote 13.

¹⁶Not everyone shares the intuition, Hinzen (2016, 25) judges *I consider John Bill* acceptable and precisely uses it as a counter-argument to the claim that there is a semantically substantive *be* with the meaning of the equality sign of mathematics.

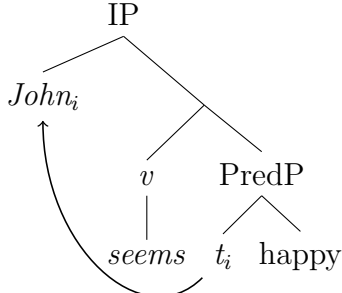
- (47)
- a. The real problem remains what to do next.
 - b. The best solution remains instant retreat.
 - c. At this point our real problem becomes John.
 - d. The critical problem now becomes how to set the parameters

Heycock and Kroch (1999, 381) therefore conclude that ‘... equative semantics is independent of the presence of the copula’.

So the conclusion from all this is that copulas don’t appear to assign theta roles.¹⁷ As such they are not like ‘real’ (lexical) verbs, and if Pereltsvaig (2001, 50) is right in the quote we gave at the start of this section, then this again supports our view of section 4.3 that the copula is no kind of lexical element.

4.4.2 The copula as a non-case-assigning raising verb

The copula is typically treated as a ‘raising verb’ (originally due to Stowell (1978)). A raising verb is a verb like *seem*, which is generally analysed (in a highly simplified way)¹⁸:

- (48)
- a. John seems happy.
 - b. 

i.e. the subject is raised from its original predicate-internal position to the subject position, [Spec,IP].¹⁹ The evidence that the subject originates below the copula comes from the same

¹⁷This is not universally agreed upon by syntacticians. See e.g. Carnie (1995) for the view that the copula in equatives assigns two theta roles, an ‘attribute’ (for the post-copula flanker) and ‘attribute recipient’ (for the subject) from which structural asymmetries between the flankers (noted prior to Carnie by Moro (1991) and Heycock (1991; 1992)) follow. Bowers (1991) also argues that the copula is a transitive verb that assigns two theta roles; see den Dikken (2006a, 71–2) for arguments against treating the copula as a transitive verb.

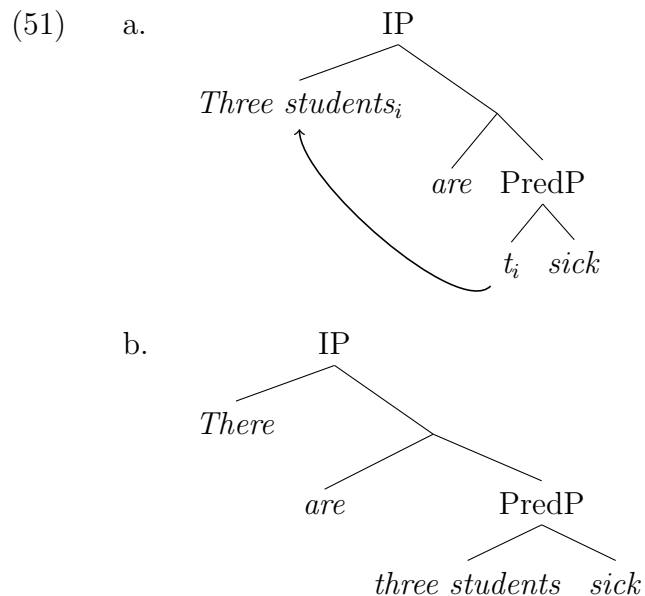
¹⁸I’m following Bowers (1993) and Roy (2013) and calling the complement of *be* PredP here. Stowell (1981) originally thought it was a ‘small clause’, a combination of subject and predicate without any intervening or embedding functional structure, in particular no tense.

¹⁹Since Larson (1988) and the ‘VP-internal subject hypothesis’, see section 2.3.2, verbs and predicates quite generally are analysed as raising constructions, e.g. transitive verbs are a complex of a VP with a light *v* (see section 2.3.4), with the subject either introduced by *v* and raising to [Spec,TP] or originating VP-internally and raising past *v* to [Spec,TP]. The difference between traditional raising verbs then and ‘normal’ transitive verbs would now just be to do with the presence or absence of a lexical VP layer embedded in the raising verb (for an explicit defence of this view see Becker (2000, 8): ‘I will argue that the main clause present

sort of facts about floating quantifiers that we encountered already in section 2.3.2 (by way of providing evidence for the VP-internal subject hypothesis):

- (49) a. All the students are sick.
 b. The students are all sick.
- (50) a. Three students are sick.
 b. There are three students sick.

The explanation for (49) is that *all the students* originates predicate-internally and then either moves in its entirety to [Spec,IP] (49a) or strands *all* (for reasons that need not concern us) in the original predicate-internal position. For (50) *three students* either raises to [Spec,IP] or stays in-situ, in which case an expletive *there* must be inserted in [Spec,IP] to satisfy the ‘extended projection principle’ (alluded to earlier, section 2.3.5), essentially the requirement that clauses have subjects. So for (50a), we have (51a), and for (50b), the structure in (51b).



Raising takes place so that the predicate’s argument receives case and is licensed. It is a standard assumption since at least Rouveret and Vergnaud (1980) that noun phrases must be ‘licensed’ by receiving case. Chomsky (1981) formulates the ‘Case filter’, which says that a noun phrase without case is ungrammatical. This explains contrasts like *For John to climb Everest would be amazing* vs. **John to climb Everest would be amazing* because *for* is taken to assign case to *John*. In Chomsky (1981), this licensing was essentially required in order

tense (finite) copula does not in fact raise from a V head, and that simple copular constructions...lack a VP’).

to make the noun phrase's theta roles legible to the interpretive systems that syntax feeds. In (50a), this raising has to take place; there is no other way for case to be assigned to *three students* if it isn't raised to [Spec,IP]. In (50b), the raising is prohibited because *there* occupies [Spec,IP] instead. How, then, is *three students* licensed? The usual solution is to say that an expletive (*there*) and its 'associate' (*three students*) form a chain (originally due to Chomsky (1981)) and that it is the entire expletive + associate chain that is assigned case, such that if the expletive is assigned nominative case, the post-verbal associate noun phrase in its chain is also licensed (Haegeman & Guéron, 1999, 243).

Notice also, that an analysis like the immediately foregoing is theoretically elegant: it unifies the assignment of theta roles into a 'predicate-internal subject hypothesis' (not just the VP-internal subject hypothesis we encountered in section 2.3.2), something which is desirable in light of the claim just above (at the start of section 4.4.1) that all lexical categories assign theta roles, not just verbs.

The importance of all of this is that it is consistent with this analysis that the copula does not assign case itself. If the copula does not assign case either (i.e. along with not assigning theta roles), then its status as a lexical verb is further diminished.

Whilst *seem* and *be* are both raising verbs, there are differences between them which are interesting to notice as they are further evidence for doubting the copula's status as a lexical verb. Whilst the subject of a sentence with *seem* like (48a) is raised out of its predicate-internal position, the verb *seem* itself cannot raise to I (the locus of the inflectional categories); *do*-insertion is required to carry the inflectional categories in negatives and questions:

- (52) a. *John seems not sad.
 b. John does not seem sad.
 c. *Seems John sad?
 d. Does John seem sad?

The copula, however, is 'light enough to be attracted by the functional heads of the I-system' (Haegeman & Guéron, 1999, 323), so the pattern of acceptability is the opposite of *seem*:

- (53) a. John is not sad.
 b. *John does not be sad.
 c. Is John sad?
 d. *Does John be sad?

The difference between the copula and *seem* for Haegeman and Guéron (1999, 323) is to say: 'To distinguish them, we propose that while copula *be* does not assign a thematic role

to its complement, *seem* is a genuine lexical verb and does assign a thematic role to its complement. Thus *seem* is not ‘light’ enough to be attracted by the inflection’.²⁰ This then is further ammunition, along with the arguments in 4.4.1 to consider that the copula does not assign theta roles.

Somewhat against this view, Safir (1985) holds that the copula is distinct from what he calls ‘identificational *be*’ (essentially a kind of identity *be*) with respect to case assignment. He believes that

(55) It’s me/I.

is ‘straightforward evidence’ that his ‘identificational *be*’ assigns case. The reading is ‘clearly identificational’ for Safir (1985, 118) in that ‘the speaker is identified’, in the same way as ‘the real villain’ is in:

(56) The real villain is him.

The point is that in both (55) and (56) the post-copula flanker has case, accusative (*me*, *him*) or nominative (*I*), and Safir believes this must be assigned by ‘identificational *be*’. And in general anywhere where a post-copula noun phrase can appear grammatically is evidence for the case-assigning role of ‘identificational *be*’ for Safir. Sentences like B’s response in (57) are judged acceptable:

(57) A: What’s left in the back room?
B: There’s only the picture of Guitar Slim.

Safir supposes here that *there* stands for ‘some discourse-controlled presupposed heading of the list’. Another assumption here is that definite noun phrases can only acceptably appear if they are assigned case. So the fact that (57b) is fine suggests that *the picture of Guitar Slim* is receiving case, which Safir says is from ‘identificational *be*’. But as we’ve just seen (just under example (51)), it seems that the way that *the picture of Guitar Slim* is licensed is by being the associate of *there* and thus forming a chain with it, a chain which is assigned nominative case via *there* being in [Spec,IP]. The parallel with (51) *There are three students sick* is even stronger if we assume that B’s reply in (57) is an elliptical form of:

²⁰Haegeman and Guéron (1999, 323) further adduce the possibility of taking a PP complement as evidence for *seem* having a thematic structure and the copula not having one:

(54) a. John seems **to me** quite happy with the solution.
b. *John is **to me** quite happy with the solution.

They don’t elaborate this any further. It seems that ‘seeming’ must be ‘of something’ and ‘to someone’ (where the latter, it seems, can be optionally spelt out), which could well be a thematic role (though which ones exactly is unclear to me).

(58) B': There's only the picture of Guitar Slim in the back room.

But I confess that I don't know how to explain the accusative case marking on examples (55) and (56). Also, if we adjusted (57) slightly, there would apparently be accusative case-marking, despite me just saying that the post-copula argument should be in a nominative case-marked chain with *there*.

(59) A: Is there anyone in the back room? Are John and Angie there, or only John?
B: There's only him/*he.

I'll leave this as problematic evidence to deal with for the view here that the copula does not assign case. One could say in reply that it is not cross-linguistically universal that the post-copula constituent is accusative in these 'identificational' cases. As Van Peteghem (1991, 18) says, '...languages like German or Russian, which mark case morphologically, always use nominative for the post-copula term in an identity'.²¹ Sandmann (1979, 211) makes the interesting observation that identity statements may really be instances where 'two nouns are coordinated first terms or subjects of an intransitive verb' which explains the nominative case-marking and 'accounts for the peculiar constructional flavour of the so-called predicative noun, which is really only a feeling of coordination as distinct from subordination'. I shan't pursue Sandmann's intuition here, though it is interesting to note that there are languages which do indeed need to express identity statements with coordinated subjects e.g. Scottish Gaelic (data from den Dikken (2006a, 75)):

- (60) a. *'s e Cicero Tully. Scottish Gaelic
COP 3sg Cicero Tully
b. 's e Cicero agus Tully an aon duine.
COP 3sg Cicero and Tully the same person
'Cicero and Tully are the same person.'

Later, in section 6.5.1, we'll see that Roy's characterisation type of predicational copular sentence can arise with at least two different syntactic structures/flanker property specifications. The example in (60) is another example of this phenomenon. The apparently same function (identity) is fulfilled in different ways in different languages; not all languages have true equatives.²²

²¹Original: ... des langues comme l'allemand ou le russe, qui marquent morphologiquement les cas, utilisent toujours le nominatif dans l'attribut d'identité.

²²Den Dikken (den Dikken, 2006a, 75–6) is a particularly neat explanation of the distribution of true equatives cross-linguistically. He analyses them as structurally the same as specificational pseudoclefts (except with the precopular flanker analysed as a reduced, i.e. null, free relative) such that languages either allow both specificational pseudoclefts and true equatives (like Welsh) or neither (like Scottish and Irish Gaelic).

4.5 Conclusion

In this chapter we have established certain key points about the copula. We saw (in section 4.2) that many languages employ nominal sentences, disposing of a copula entirely in many contexts. This in itself is suggestive of the way the copula is hardly a semantically substantive constituent in a sentence; semantically substantive constituents need to be present in a sentence in order for their meanings to be included as part of the compositional meaning of a sentence. Copulas are obligatorily present for many reasons (person, tense etc.) and whilst there is some apparent evidence that copula obligatoriness correlates with an equative reading 4.2.3, this is not cross-linguistically universal (see the examples in (42)) and the prediction of this thesis is that these facts must be explained in some other way (though admittedly I do not provide such an explanation). We then gave some quite mainstream linguistic analysis of the copula which allowed us to conclude that the copula is not a lexical, but rather a functional, category, section 4.3. Furthermore, it isn't even really typically verb-like, section 4.4. For example, the copula does not assign a theta role to a subject; rather this is assigned directly by the predicate in a copular sentence. In that respect, it is unlike lexical verbs which—by definition—have a theta grid. In case-marking terms, the copula is similar to an unaccusative verb; it is a verb with one (small clause) argument that it does not assign case to, and so which needs to raise to [Spec,IP] in order to be case-marked with nominative case (*unlike* unaccusative verbs, the copula does not assign a theta role to its argument though, as we argued in 4.4.1).

Here is a nice final word on the issue of whether the copula is some kind of truly lexically ambiguous element, ambiguous between predication and identity:

Assuming the existence of two verbs *be*, rather than two (or more) uses of the one verb, leaves unaccounted for the fact that the two verbs are phonetically identical. While this could well be an accident in English, most of the languages I have examined have only one verb for the two uses. If there were indeed two verbs, the fact that in (almost) every language the two verbs are phonetically identical would be an extremely odd, and inexplicably universal, coincidence. (Rapoport, 1987, 142).

It seems completely improbable that a genuine lexical ambiguity would be pronounced the same way cross-linguistically. Languages don't work like that; such putative 'cross-linguistic homonyms' do not exist.

Chapter 5

What the copula does: two case studies

5.1 Introduction

So we can conclude from chapter 4 that the copula is not a typical semantically substantive verb. This is all evidence to suggest that it isn't capable of being meaningful enough to act as a semantically ambiguous constituent which can distinguish the meanings of different types of copular sentence. But it would be wrong to conclude that the copula does *nothing* with respect to the meanings of the sentences that it appears in. But what it does exactly is not cross-linguistically uniform, so in this chapter I shall simply present two case studies, from English 5.2 and Spanish 5.3, to give an idea of at least two ways the copula is operating in two different languages. Making a general conclusion from such a small sample of the world's languages is impossible, so the conclusions will be somewhat limited. As I said in the introduction to chapter 4, my most general view is the rather unsatisfying conclusion that the copula is a 'sign of structure'. This accords with current research, e.g. Arche *et al.* (forthcoming) say, '...copulas lack 'conceptual' content, but contribute structural semantics... [the] copula is not a distinct grammatical category, but rather the label that has been given to a number of distinct objects in different languages... what we call 'copula' in one language is quite likely different from what we call 'copula' in another; copulas seem to be involved in different syntactic constructions cross-linguistically' (Arche *et al.*, forthcoming). In some languages, they are essentially analogous to focus markers (Faarlund, 2007; Zellou, 2010; Hartmann & Veenstra, 2013), in others they are devices used to introduce argument structure (Welch, forthcoming), and yet others they are used 'to define passive voice' (Arche *et al.*, forthcoming), amongst much else. Below, we shall see two different ways that the

copula is used.

5.2 The copula in English: a case study

As we noted in chapter 4, nominal sentences—that is, sentences in which a copula is not present, such as Hungarian *Ő diák.*, literally ‘He a student’ meaning ‘He is a student.’—do not exist in English. However there is one occasion where the copula can go missing, in sentences like:

- (1) a. John considers Mary to be intelligent. *infinitival*
 b. John considers Mary intelligent. *small clause*¹

i.e. where the erstwhile ‘copular sentence’ is embedded under certain verbs like *consider*. Both (1a,b) are grammatical and seemingly semantically equivalent and interchangeable; one could apparently use either (1a) or (1b) on any given occasion without there being any obvious reason why one would use one and not the other, an apparent case of free variation. As such, it provides a straightforward piece of evidence for anyone who wants to claim that the copula is semantically vacuous. If it wasn’t, then there should surely be some difference between (1a) and (1b). This was the basic point we made in section 4.2. The point of this section, therefore, is to check whether this is really the case. The conclusion will be that there are in fact sentences like (1a,b) where the presence of the copula makes a seemingly semantic difference—sentences where the subject of the sentence as a whole—the ‘matrix’ subject (*John* in (1))—and the subject of the small clause/infinitival—the embedded subject (*Mary* in (1))—are quantified terms (see (5) below). We conclude from these sentences that (1a) is structurally different to (1b), a difference that shows up in different meanings with quantified terms, in terms of different possible scope relations. Where the matrix and embedded subjects are just singular terms, however, this particular structural difference has no effect; in this context (with the singular terms), then, it is a case of a structural difference which makes no semantic difference. It would therefore perhaps be more accurate to say that (1a) and (1b) are distinct syntactic structures that happen to be ‘semantically interchangeable’ when the matrix and embedded subject are both singular terms; this is perhaps a better way of putting it than saying that the copula can ever freely be present or absent in (1a,b). The copula is a reflex of the structure of (1a). It *has* to be there when the structure underlying (1a) is used in English, and it is precisely what a hearer uses to understand the underlying structure of the sentence used. It just so happens that whether

¹Reminder: a small clause is a combination of a subject and predicate without any intervening or embedding functional structure, in particular no tense.

this structure is used or the one in (1b), the semantic effect is the same when the matrix and embedded subjects are singular terms (but not when they are quantified terms).

So now I need to explicate what I've just said. Firstly, what is the evidence for different meanings for (1a,b) when the matrix and embedded subjects are quantified terms, and secondly what is the syntactic analysis which might explain these different meanings? Both these questions will be addressed in turn in the following two sections.

5.2.1 The evidence for different meanings when the matrix and embedded subjects are quantified terms

Unlike the examples I gave in (1), the first set of examples I'll give relate to the scope of another embedding verb *seem*, not *consider*. *Seem* shows the kind of difference we'll be interested in when we replace the subjects of (1) with quantified terms, namely a scope difference whereby the matrix subject either *has* to have wide scope (in the small clause complement case), or where there is an ambiguity between wide or narrow scope for the matrix subject (in the infinitival case). The difference between the examples with *seem* and examples with verbs like *consider* is that there is no distinct embedded subject of the small clause or infinitival, so we can't detect different scope relations between the two subjects (the matrix and embedded one). However, *seem* is a verb which, itself, has a scope domain; anything in the scope domain of *seem* merely seems to be the case, whereas anything outside its scope domain is rather asserted or presupposed to be the case by the speaker. The examples below should make this pre-amble clearer. I'm starting with these *seem* examples because I think the scope difference judgements are clearer than the sorts of examples I gave in (1).

Having said this, the original examples from the literature are not, frankly, very clear judgements, so I shall just cite these examples, as they are given, before giving some more recent examples that I think make the judgements clearer. One of the earliest examples of the type we're interested in comes from Williams (1983a, 293), who gives these examples (cited in den Dikken (2008, 1) with the scope relations indicated²):

²' $\exists(a \text{ student}) > \text{seem}$ ' means that *a student* (or more properly, the existential operator in the logical representation of e.g. (2b)) takes wide scope with respect to the embedding verb *seem*. If we think of a word like *seems* as having a scope domain whereby anything within that domain is modified by *seems*—i.e. whatever it is seems to be the case, rather than being asserted or presupposed to be the case—then we would say that any constituent that takes wide scope with respect to it is not so modified. This is in contrast to (2a), where $\text{seem} > \exists(a \text{ student})$ is a possibility i.e. where *a student* can take narrow scope with respect to *seems*, and so the existential operator in the logical representation of (2a) merely seems to be the case, unlike (2b) where the existence of a student is asserted. So (2a) isn't falsified if it turns out that whatever seems to be sick is not in fact a student, nor if whatever seems to be a student turns out not to be sick; nor indeed if it turns out that there is neither a student nor a sick person, but merely a ventilation shaft making its

- (2) a. A student seems to be sick (ambiguous: $\exists(a\ student) > seem; seem > \exists(a\ student)$)
- b. A student seems sick $\exists(a\ student) > seem; *seem > \exists(a\ student)$

As Williams puts it:

- (3) a. A student seems sick.
- b. = There is a student who seems sick.
- c. \neq There seems to be a student who is sick.

To perhaps render this judgement slightly clearer, Moulton (2013) gives the following context to (2): ‘Seeing an empty seat in our otherwise crowded classroom’. The idea with this is that (2a) will be still be acceptable, but (2b) won’t be. (2b) requires (as per ‘There is a student who seems sick’) that there actually be a student in the room who we could in theory point to. (2a) doesn’t necessarily assert the existence of a student, and so the reading where *a student* takes narrow scope relative to *seem* would be an acceptable utterance in Moulton’s context.

If the judgement here is still not completely sharp, then Moulton gives another example, which I think is the clearest we’ll get:

- (4) Context: Last weekend we were walking through the woods and encountered a sleeping bear. But it was so large we mistook things.³
- a. Two bears seemed asleep. $2 > seem; *seem > 2$
- b. Two bears seemed to be asleep. $2 > seem; seem > 2$

As Moulton (2013, §8) puts it: ‘A wide scope interpretation for the raised subject requires that there are two bears and given the available evidence those bears are asleep. The narrow: given the available evidence, there are two bears asleep. In the scenario, only the narrow scope construal is true’. This is because the context explicitly tells us there is ‘a sleeping bear’ i.e. only one bear is presupposed. As such we couldn’t use (4a), as it presupposes that there are two bears, in direct contradiction with the presupposition given in the context. (4b), on the other hand, has an available reading that does not make this presupposition, and which is therefore felicitous, following on from the given context.

Whilst the clearest examples involve *seem*, essentially the same scope difference is also reported by Hornstein (1995, 76) in the following examples, using *consider*, like the first examples we gave in (1):

usual noise. (2b), on the other hand, is falsified if there turns out to be no student (but not if it turns out that the student is in fact not sick). Here, I shall talk of *X* ‘scoping over’ *Y* where $X > Y$.

³By ‘mistook things’, I take Moulton to mean that we were mistaken about what we thought we saw.

- (5) a. At least one person considers every senator to be smart.
 b. At least one person considers every senator smart.

As Hornstein (1995, 75) puts it: ‘It appears that the universal QNP [quantified noun phrase, i.e. *every senator* in (5)] in (5a) can have scope over the existential NP but this is virtually impossible in [(5b)]’. The idea is that in (5a), there is the possibility of a ‘distributive’ reading whereby for every senator there may be a different person who considers them smart (i.e. John considers senator Smith smart, Mary considers senator Jones smart etc.). In (5b), on the other hand, there can only be one same person who has a thought about all the senators that they are smart (i.e. John considers senators Smith, Jones etc. smart). This is why den Dikken (2008, 6) gives this judgement:

- (6) A (*different) student considers every congressman a fool.

Hornstein (1995, 213, n.8) further says: ‘Observe that [(5a)] can be augmented with the phrase “namely his mother” forcing the wide scope universal reading. This is not possible in [(5b)]’. The idea here is that ‘his mother’ picks out each individual senator’s mother as the antecedent of ‘at least one person’; this is only possible if *every senator* takes scope over *at least one person*, such that there is a different person (a different mother, say) assigned to each different senator. Of course, there is a possible reading of (5b) *At least one person considers every senator smart, namely his mother* where *his* is a referential pronoun (rather than a bound variable one as it could be after (5a)) i.e. a salient person, John say, whose mother considers every senator smart. Hornstein’s point is that this is the only reading we could get for (5b); we couldn’t get the bound variable reading as we can in (5a).

5.2.2 Syntactic analysis of the different scope relations—‘small clause restructuring’

So we are interested now in accounting for the scope differences in sentences like (5). The way that accounts within Universal Grammar have worked is essentially based on the idea that in sentences like (6), the small clause predicate *smart* incorporates with *considers* to form a complex predicate *considers+smart*, a process called ‘small clause restructuring’. Here we will focus on the most comprehensive attempt at an account of these sentences, (den Dikken, 2008), after running through some background to this account in Stowell (1991) and Hornstein (1995).

5.2.2.1 Stowell

Stowell (1991) appears to be the first to use the idea of small clause restructuring, though Stowell himself cites Chomsky as the originator of the basic idea. Chomsky's own analysis was essentially the opposite of Stowell's, however, in terms of the mechanics of the operation. Rather than starting with the small clause as a unit, the predicate of which joined with the matrix embedding verb later, Chomsky (1975) had the complex predicate, e.g. *consider+smart* formed in the first place at 'D(eep)-structure' before a transformational rule moved *smart* to the right at S(urface)-structure. The difference comes from the way that a different model of UG was in use at the time, notably the difference in terms of where the interface between syntax and semantics was located. For Chomsky, this was D-structure; subsequent syntactic transformations to produce S-structure did not feed back into semantics. Later on, D-structure was eliminated in favour of a model in which we simply start with a bundle of lexical primitives, extracted from the lexicon (the long-term storage site of lexical items); these are then constructed into 'syntactic objects' (to take Chomsky (1995)'s term) which are subsequently fed into the semantic interface ('Logical Form'—LF), as well as the interface with the system(s) for pronunciation (of oral languages).

Whilst small clause restructuring will be useful later for accounting for sentences with two quantifier phrases—QPs—in matrix and embedded subject position, for Stowell it was only used to account for scope relations with verbs like *seem*, i.e. the difference between (7a,b) that we looked at earlier:

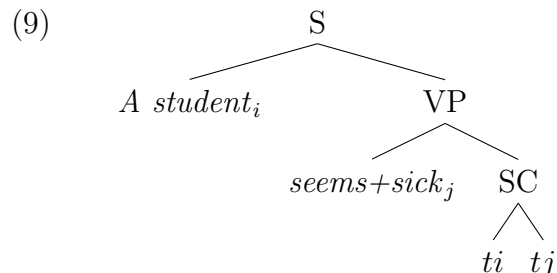
- (7) a. A student seems sick.
 (i) = There is a student who seems sick.
 (ii) ≠ There seems to be a student who is sick.
 b. A student seems to be sick.
 (i) = There is a student who seems sick.
 (ii) = There seems to be a student who is sick.

Stowell's view is that once the small clause predicate has moved to incorporate with the matrix verb, '[t]he subject [i.e. the matrix subject *a student*] then cannot lower because there is no longer a predicate within the SC [small clause] for the quantified NP to take scope over'. To unpack this, it is based on the following principle:

- (8) *Predicate Scope Principle*
 a. A quantifier phrase QP must take scope over a predicate P.
 b. For any predicate head P appearing in a chain of linked positions (P, t_1, \dots, t_n), QP takes scope over P if and only if QP c-commands P (the head of a chain

consisting of a moved predicate and its traces i.e. the previous positions it has occupied.)

So for *A student seems sick*, we would get a tree as follows:



The raised *a student* would not be able to lower to adjoin to SC (and so be in the scope of *seems*) because then it would not c-command the head of the chain of the raised predicate *sick*, in violation of the principle in (8).⁴

So to re-state the logic of Stowell's argument: the stipulation is that any quantifier phrase just must take scope over some predicate. To do this, it must c-command a predicate. So in the tree in (9), we can see that *a student* c-commands the complex predicate *seems+sick*. This means (according to the principle in (8)) that it can take scope over a predicate, and so there is a wide scope reading of *a student*. For the narrow scope reading of *a student* to be possible, we would need to lower it to its trace position (marked t_i in the tree in (9)). However, because the predicate *sick* has moved, *a student* would no longer c-command it from this lowered position (and clearly it doesn't c-command *seems* either). It would c-command the trace of *sick*—marked t_j in (9)—but this is not the head P of the chain (P, t_j) and as such (again according to (8)) the QP cannot take scope over P. And given the stipulation that a QP must take scope over a predicate (and this is the case even here where we precisely want the QP *a student* to be in the scope of *seems*, and are not looking for a scope relation between *a student* and *sick*), then the lowering of *a student* is simply an infringement of this principle, and therefore an impossible move. It isn't clear to me exactly what this principle is, but it seems most likely that it was intended as a rule about possible Logical Form (LF) movements i.e. covert movements that feed the LF interface

⁴Basilico (2003, 25) notes a problem for this approach, initially observed by Williams (1983a); the prediction Stowell makes is that arguments of small clauses where the small clause predicate incorporates with the matrix verb will never be able to take narrow scope in relation to the matrix verb. However, in (i), it seems *something* is in fact within the scope of the matrix verb:

(i) He seems proud of something.

The judgement—which seems fair—is that there is a reading here where we aren't necessarily presupposing that there is something which he seems proud of. Having said that, it strikes me that even in *He seems to be proud of something* (which we shall soon see should allow for a wide scope reading of *something*), we hardly get a reading where there is some particular thing that he seems proud of.

between syntax and semantics. It should perhaps be pointed out here that the positions that quantifier phrases take scope from are always associated with their LF position, and any ‘movement’ that occurs in order to put a given quantifier phrase in a given position at LF (in order to account for a particular scope interpretation) is LF-movement, which is always covert movement. Sometimes this coincides with the surface level position of the quantifier, sometimes not.

5.2.2.2 Hornstein

Hornstein (1995, 247) gives a partial explanation of the different scope possibilities in sentences like (5), repeated here with the examples Hornstein gives there:

- (10) a. Someone considers every congressman a fool.
 b. Someone considers every congressman to be a fool.

In (10b), the sentence is ambiguous between wide scope readings for either *someone* or *every congressman*. Hornstein seems to assume that the wide scope reading for *someone* is not in need of explanation. As just mentioned, in syntax a constituent can scope over another if it c-commands the other constituent. Presumably, for Hornstein, the fact that in its surface position *someone* is visibly higher up the syntactic tree than *every congressman* in (10) justifies the assumption that it could be c-commanding it and so there is no problem deriving the the wide scope reading for *someone*.

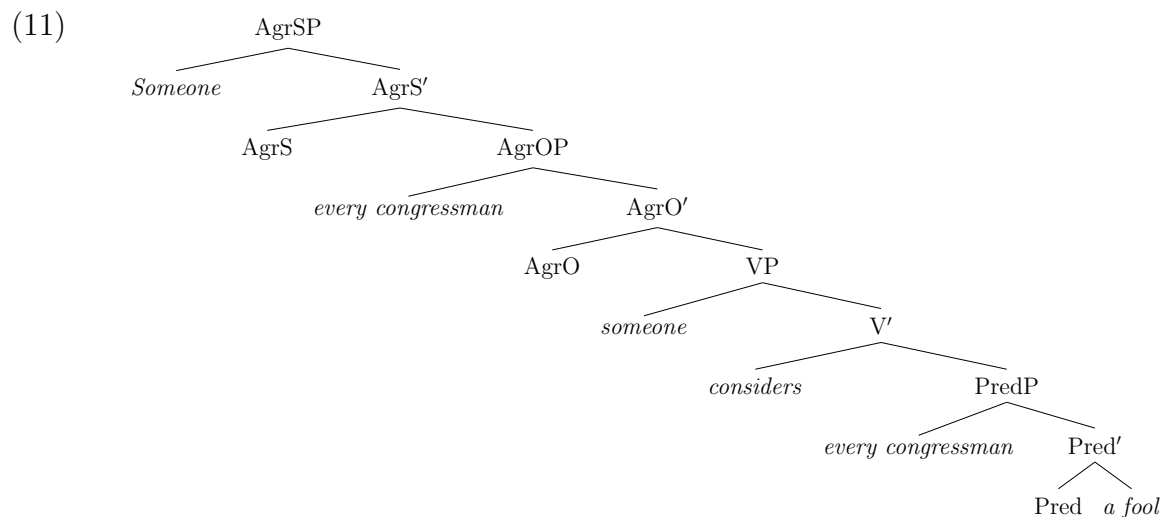
Hornstein then says that the ability to derive a wide scope reading for *every congressman*, ‘is what we expect given that in these small clauses *every congressman* moves to the matrix Spec,AgrO’.⁵

So Hornstein assumes that we should expect *every congressman* to move to the specifier of AgrO such that it should be able to take scope over a ‘reconstructed subject’ (247). This is visible in the tree below, (11). A ‘reconstructed subject’ here basically just means an interpretation of the sentence where the subject (in this case *someone*) is interpreted in one of its lower positions in the tree i.e. in the tree of (10b) in (11) below, the reconstructed

⁵‘AgrO’ was a functional head—the locus of object agreement features and part of the extended projection of the verb—that featured prominently in the ‘Principles and Parameters’ theory of UG (Chomsky & Lasnik, 1993; Belletti, 2001), along with AgrS—a subject agreement phrase. Arguments of the verb moved to the specifier of AgrO and AgrS in order to be licensed (via case-marking) as objects or subjects respectively. The postulation of these distinct functional heads made sense of some agreement patterns (e.g. word order facts) cross-linguistically. These functional heads were subsequently subsumed by a theory of movement driven by uninterpretable ‘phi-features’ (number, person, gender features) which are properties of lexical items that enter the derivation, and that are checked and eliminated in a probe-goal relationship with a counterpart lexical item holding interpretable versions of the same features, and without the necessity for separate functional projections (originally proposed in Chomsky (1995, ch.4), and then developed in Chomsky (2000; 2001)).

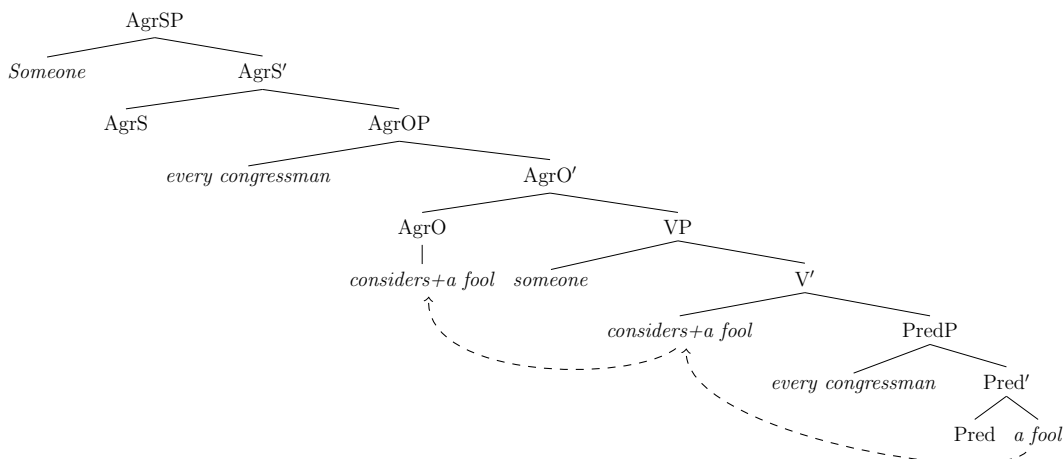
position of *someone* is the specifier of VP where it originates and from which it moves to the specifier of AgrS.

Here then is a tree which shows the possible scope positions that the two quantifier phrases could theoretically occupy, and which would give rise to an ambiguous sentence of English if it were its underlying structure; the ambiguity would be resolved according to which positions the quantifier phrases occupy at LF. Presumably, Hornstein thinks that something like this is the right analysis of (10b), the infinitival case, which is ambiguous. But the infinitival copula is not shown here:



For Hornstein, then, the default situation is that *someone* should be able to ‘reconstruct’ to the Spec,VP position where it originates, from where it would be scoped over by *every congressman* in the Spec,AgrOP position. So, for Hornstein, the question is why this is not possible in (10a). Hornstein’s solution rests on the idea that predicates ‘like all other expressions’ must be licensed. Furthermore, Hornstein assumes that the licensing of predicates is done by incorporation into/moving to the sorts of functional heads that usually go with predicates, such as AgrO. So a typical verb (like the *consider* of (11)) was analysed at the time Hornstein was writing as raising to AgrO as part of its licensing requirement (this movement isn’t actually shown in (11), but we see the functional layers AgrOP and AgrSP above VP). However, small clauses have no layers of functional structure above them. How then does the predicate of a small clause get licensed if it can’t simply rise to an AgrO within the small clause? This is where Hornstein uses Stowell’s small clause restructuring; he says that the predicate incorporates into the matrix verb. The structure that we get then get for (10a)—the small clause case—is:

(12)



The licensing that occurs here is this: (1) The predicate *a fool* incorporates with the matrix verb *consider*, this whole complex then rises to adjoin with AgrO. (2) *every congressman* is licensed by raising to Spec,AgrOP. (3) *someone* is licensed by raising to Spec,AgrSP. So everything is licensed. Sadly though, a real solution to the problem (of the different possible scope relations in (10)) is lacking: Hornstein (1995) merely says that the incorporation of *a fool* into *consider* ‘...somehow prevents the reconstruction [i.e. of *someone* to the Spec,VP position below *every congressman* in Spec,AgrOP]. I have no idea why this would be so at the moment...I leave it as an exercise for the reader to come up with a reason why predicate incorporation prevents reconstruction’. So now let’s turn to a reader who took up that challenge.

5.2.2.3 Den Dikken

5.2.2.3.1 The rationale for small clause restructuring Den Dikken (2008) essentially agrees with Hornstein about the ultimate rationale for small clause restructuring, the process that incorporates the small clause predicate with the matrix verb. Without doing this, the small clause predicate is not licensed. Specifically, den Dikken claims that small clauses need to be licensed by a temporal or aspectual head; this is unlike Hornstein for whom the licensing was via the predicate raising to AgrO, though den Dikken is writing over ten years later when categories like AgrO were not so typical in minimalism; agreement is now generally accounted for in terms of bundles of agreement features that can be situated on lexical categories themselves or functional phrase heads like Tense and Aspect. In the case of the infinitivals, the small clause is the complement of Tense (as we see in (18) below) and is thus licensed by a Tense head. But in the small clause case, the only way for the small clause to get licensed by Tense or Aspect is if the head of the small clause (marked ‘R’ for ‘Relator’ in (18) below) incorporates into ‘the T[ense]-chain of the verb’ (den Dikken, 2008).

5.2.2.3.2 Basic premises of the account The essential point of den Dikken (2008)'s analysis is that (at least part of) the job of syntax is to describe scope domain sizes, which den Dikken aligns with the 'phase'. What is a 'phase'?

A consistent notion in the history of Universal Grammar is that of an interval which demarcates an upper limit for various linguistic generalisations or operations and which thereby reflects the locality of syntax. This has gone by different names over the years, such as 'cycle', 'barrier', or 'bounding node' (Gallego, 2010, 51), and now is called a 'phase' (Chomsky, 2001; Chomsky, 2013). There is widespread agreement that something like a phase exists, but widespread disagreement about what a phase is, where they come from, and whether particular kinds of phase are cross-linguistically universal. To render this slightly more concrete, an example of the kind of thing to which locality applies is the form of pronouns. In the case of English there appears to be complementary distribution between reflexive and non-reflexive pronouns, such that in *Peter said John shaved himself*, the reflexive is only co-construed with *John*, whereas in *Peter said John shaved him*, the non-reflexive pronoun is only co-construed with *Peter*. In a sense which linguists have sought to make precise, *Peter* is too far away for the reflexive and *John* is too close for the non-reflexive pronoun. The size of the relevant local interval could be roughly characterized as the clause, though the literature on this question (known as 'binding theory' (Safir, 2004a)) is substantially more nuanced than that. Having said that C, the phrasal category corresponding to the clause, is one of the more uncontroversially accepted instances of a phase (Citko, 2014) (see Hicks (2006) for a phase-based implementation of binding theory). More controversy arises at intervals which are smaller than C.

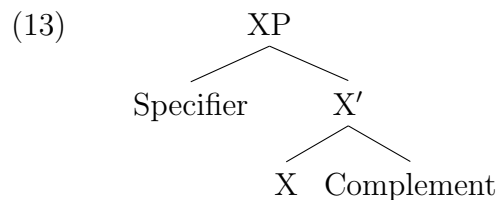
So phases are small selections of items from the lexicon (the long-term memory store of lexical items i.e words); if they were too big, there would be a problem of computational load for humans. Chomsky makes an analogy for the way that such computational efficiency may be operative in language:

The obvious proposal is that derivations make a one-time selection of a lexical array LA from Lex [the lexicon], then map LA to expressions, dispensing with further access to Lex. That simplifies computation... If the derivation accesses the lexicon at every point, it must carry along this huge beast, rather like cars that constantly have to replenish their fuel supply. (Chomsky, 2000, 100–1)

Phases are generally thought of as the smallest selections that go into any derivation, and it is a matter of debate about how many levels of selection there are, and what their respective sizes are, though something like a two-step selection of all the items one might

need for something as big as a clause, say, and then a selection from this of items for each phase in a derivation, seems to be operative in Chomsky's work.

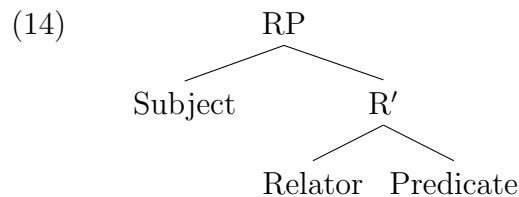
For den Dikken, one way at least of defining a phase is to say that predications are phases i.e. anything which is a predication will by definition be a phase i.e. a phrase level where local relations (like that between a reflexive pronoun and its antecedent) hold, and beyond which they don't. The next thing to say is that scope positions (and this is also a quite uncontroversial move) are phase edge positions. What is a 'phase edge'? Well, under the X'-theoretic model that we've been working with throughout this thesis, every head projects a complement and a specifier:



The 'X' in (13) is the head, the 'Complement' and anything it dominates is the interior of the phase, and the specifier and anything it dominates (and there can be more than one specifier, (Chomsky, 1995, 354)) is the edge of the phase. So QPs need to be in specifier positions of phase heads, like C (which we could plug in for X in (13)).

A novelty of den Dikken's approach is what he calls 'phase extension'. This occurs when a phase head moves to a higher head position. The effect of this is that the head that the phase head moves to itself becomes a phase head and this extends the interior of the original phase. The original phase edge ceases to be a phase edge and the edge of the head which has been moved to now becomes a phase edge. We'll see this in action shortly.

So den Dikken's version of Stowell's small clause restructuring runs as follows: the core of predication is conceived of by den Dikken as a 'Relator Phrase' (see den Dikken (2006a) for a full working-out of this approach); different phrasal heads can perform the role of a relator (e.g. verbs can be relators in verbal predications). Small clauses are not actually 'small' for den Dikken, but rather contain a null relator as follows:



LF movement (covert syntactic movement that feeds LF) raises the relator up to the verb that selects the small clause e.g. *consider* in *We consider John a fool*. This raising (see

(16) below) of the relator (a phase head) extends the RP phase up to the level of the verb (*consider*); the original specifier of RP is no longer a phase edge (and therefore no longer a scope position) and the edge of the verb (*consider*) becomes a phase edge and scope position.

Den Dikken (2008) does in fact give an argument to justify saying that there is LF movement of the relator here: there are cases where it seems there must be a spelt out relator:

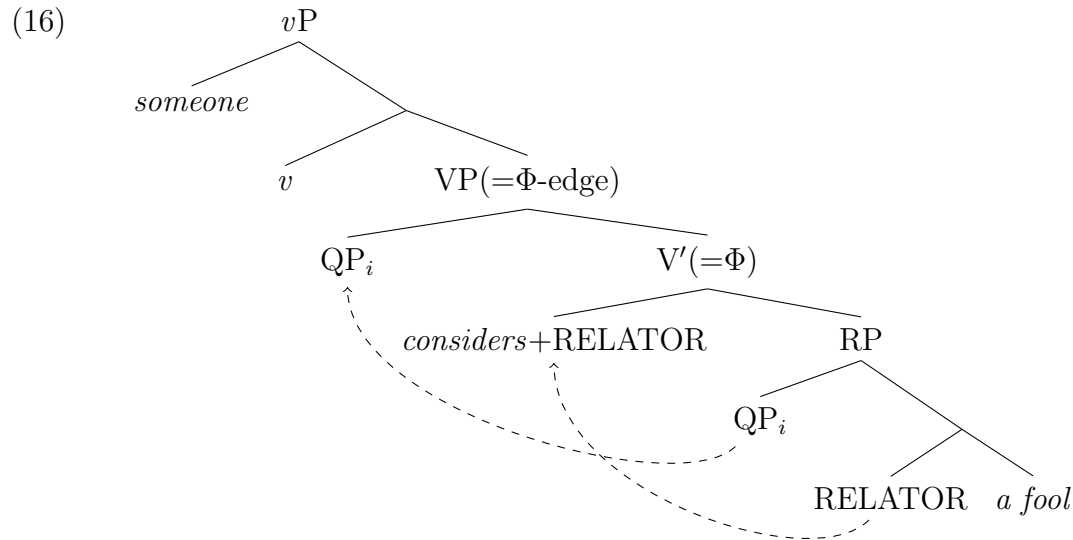
- (15) a. Someone regards every congressman *(as) a fool.
 b. Someone takes every congressman *(for) a fool.
 c. Someone makes every congressman *(out) a fool.⁶

i.e. where matrix verbs like *regard*, *take*, or *make* are used rather than a verb like *consider*. In these examples, we don't see that the relator moves when the complement of the matrix verb is a small clause. Notice crucially that the scope relations in these sentences in (15) are the same as (5b), i.e. there is no interpretation where *every congressman* scopes over *someone*; this is what tells us that 'small clause restructuring' is at work in these examples. So given that fact, and the fact that the relator (*as*, *for*, and *out*) is spelt out in its base position in the examples in (15) (i.e. the sentence isn't e.g. **Someone regards as every congressman a fool*), we can conclude that the movement of the relator to the verb required to perform small clause restructuring is a covert LF one.

So let's now run through the examples we started all this with and show how they work in the light of this set-up.

5.2.2.3.3 The small clause case—*Someone considers every congressman a fool*
 The tree for *Someone considers every congressman a fool*, then, looks like this:

⁶There may be some speaker variation with the judgement here; an American English speaker reports that this is not good as an embedded small clause i.e. it would need to be *Someone makes every congressman out to be a fool*, where, furthermore, *out* is not optional. My own (British English) judgement is that this example should be *Someone makes every congressman out (as) a fool*. My feeling is that *make out* is a phrasal verb, where *out* is not optional, and where the relator is, as in (15a), *as*.



What do we see here? Well, we see that the relator moves to incorporate with the verb *considers*. This has the effect of extending the RP phase to VP, phasehood marked here with a Φ . This makes the specifier of VP a phase edge and hence a scope position, and means that the original phase edge, the specifier of RP is no longer a phase edge and so no longer a scope position. The QP (*every congressman*) moves to the edge of this extended phase, to the specifier of VP (this must be LF movement as *every congressman* is spelt out lower than *considers*—this of course is fine because, as we’ve said, movement for scope is always LF movement). As we’ve said, phase edge positions are the positions which in this account are the ones that scope operators, like QP, must occupy. On the basis of this tree, we can see no position in which *someone* would be in the scope of the QP *every congressman*, the original observation that we wished to explain.

Why can’t the QP *every congressman* in (16) move further up the tree to a scope position above *someone*, as it does in the infinitival case (as we’ll see below in (18))? This of course is a crucial matter, but den Dikken’s (2008)’s answer to this is not completely clear to me; I’ll try to summarise the reasoning. He starts by noticing that ‘a well-known fact’ is that ‘quantifier phrases embedded in a finite complement clause (which inevitably is or dominates a phase boundary) cannot gain scope over a constituent in the matrix clause’. No example is given, but presumably this sentence should not allow a distributive reading where *every teacher* scopes over *two students*:

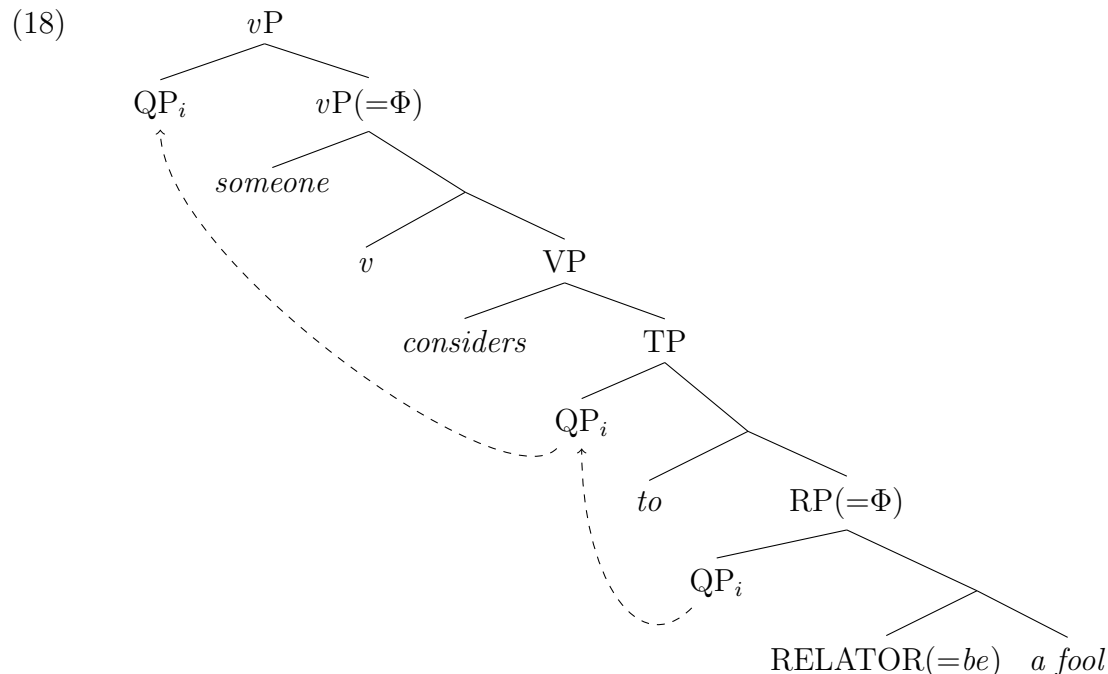
(17) Two students believe John likes every teacher.

Then he says that this fact shows that there is no step-by-step movement of quantifier phrases from one phase edge position to another. If we look ahead a little to the tree in the infinitival

case (18) below, the step-by-step movement of the QP is mediated by a non-phase edge Spec,TP position. So quantifier raising is either one fell-swoop movement between phase edge/scope positions, or step-by-step via intermediate non-phase edge/scope positions.

The key point seems to be this: the nature of quantifier raising (and LF movement generally) is that it feeds semantics, the interpretation of the sentence. As such, if a quantifier occupied multiple scope positions on its way to any destination, it would send conflicting interpretations to semantics. A quantifier would be in two positions with potentially contradicting scope relations with other constituents in the sentence. So quantifier movement must always be done in one fell-swoop. Now, den Dikken says, ‘[s]ince fell-swoop movement out of an embedded finite clause is standardly taken to be impossible. . . and successive cyclic [i.e. step-by-step] Q[uantifier] R[aising] is blocked as well, QPs embedded in a finite complement clause cannot scope over QPs in the matrix clause’. I have a couple of issues here: it isn’t obvious that small clauses are finite complement clauses in such a way that the ban on fell-swoop movement out of finite complement clauses is relevant here. The idea that successive-cyclic quantifier raising is banned (regardless of whether this is out of a finite clause or not) does ostensibly help us, but only if we make the assumption that the QP in (16) would necessarily have to be raised via the intermediate scope position (the specifier of VP, which has become a phase via phase extension). Without that assumption, and without the assumption that small clauses are finite embedded complement clauses which nothing can be quantifier raised from, then it isn’t clear why the QP in (16) couldn’t move in one fell swoop to a position above the matrix subject. Also, I think den Dikken means to say that successive-cyclic quantifier raising *via scope positions* is banned: in his own structural description of (18) below, we have successive cyclic raising of the QP via TP (a non-phase edge position). So, sadly, an entirely clear understanding of why *every congressman* cannot move into a position above the matrix subject position in (16) eludes me.

5.2.2.3.4 The infinitival case—*Someone considers every congressman to be a fool* Here, then, is the structure for the infinitival case *Someone considers every congressman to be a fool*:



We can see that phase extension via small clause restructuring is not applying to RP (it doesn't have to for RP to be licensed because the TP above RP licenses it), so RP is a phase, and its specifier is a potential scope position, i.e. a position in which the QP could be scoped over. This gives us the reading where the QP *everyone* takes narrow scope with respect to (is scoped over by) *someone*. There is no intermediate phase position because of there being no small clause restructuring. Furthermore, there is a position in the edge of vP , which is a phase, where the QP can move to, and from where it can scope over *someone*.

Turning to the tricky issue (see the end of the previous section, 5.2.2.3.3) of why there is a position here (where there wasn't in the small clause case) above the matrix subject that *every congressman* can be raised to, in order to take scope over *someone* at LF and give us the distributive reading that is unavailable in the small clause cases: the only thing that I can think to say here is that it is fairly obvious that the embedded clause *every congressman to be a fool* is not a finite clause complement. As such, QR out of it isn't standardly considered to be impossible. But as I said before, small clauses are not obviously finite clauses either. It seems like den Dikken wants to say that the small clause restructuring creates a phase that *every congressman* can't escape in the small clause case. But this phase is simply the extension of the RP, the core predication which is the very same RP that appears in the infinitival case as well. Banning phase edge to phase edge movement for quantifiers, as den Dikken explicitly does (den Dikken, 2008, 3) would have the effect of banning quantifier raising in the infinitival case too it seems to me, which would be contrary to the facts that we want to explain.

- (22) a. Esta solución está inteligente. Spanish
 this solution ESTAR.3sg intelligent
 ‘This solution is intelligent.’
 b. Ana está muerta.
 Ana ESTAR.3sg dead
 ‘Ana is dead.’
- (23) a. Su mirada es ausente. Spanish
 his look SER.3sg absent
 ‘He has an absent face.’
 b. Pedro es joven.
 Pedro SER.3sg young
 ‘Pedro is young’.

As Roy (2013, 143) puts it: ‘The general situation seems to be that an adjective can be constructed either with *ser* or *estar*, and that their distribution relates to uses rather than to lexical semantic restrictions’. In other words, it isn’t the nature of the predicate that enforces a particular type of copula. For some predicates, the permanent vs. temporary nature of a property is not differentiated between an occurrence with *ser* or *estar*. In the (24), ‘cold’ is a permanent property of the subject *La nieve* in both cases:

- (24) a. La nieve es fría. Spanish
 the snow SER.3sg cold
 ‘Snow is cold.’
 b. La nieve está fría.
 the snow ESTAR.3sg cold
 ‘The snow is cold.’

The difference, rather, comes from the way that (24a) is a general statement about snow, which could be used to teach a child what ‘cold’ is. (24b) must report the perception of the speaker, whilst they are touching some snow, for example. As Roy (2013, 144) says, ‘... the choice of the copula relates to the speaker’s perspective on the relation between the denotation of subject and a particular property or a particular state’.

The following example has a predicate that appears to denote a temporary property, the transience of the property rendered explicit by the temporal delimitation; we could therefore expect it to be acceptable with *estar*. However, as the acceptability judgement indicates, the sentence is not good:

- (25) a. *Juan está profesor (en sus horas libres). Spanish
 Juan ESTAR.3sg professor in his hours free
 intended: ‘Juan is a professor (in his spare time).’

Table 5.1: Distribution of auxiliaries and predicate form and interpretation in Spanish, Roy (2013, ch.6)

	dense	non-dense	maximal
copula	<i>estar</i>	<i>ser</i>	<i>ser</i>
predicate	AP, PP	NP, Adj/PP-N _{pro}	NP, Adj/PP-N _{pro}
		no article	article

In line with her overall approach to ‘predicational’ copular sentences (essentially, for Roy, all copular sentences except those that have a DP predicate), Roy rather considers there to be a three-way distinction as in Table 5.1.

This, then, is a purely ‘categorical’ approach to the distribution of *ser* and *estar*; the essential idea is that *ser* goes with nominal predicates, and *estar* goes with all non-nominal predicates.

A potential problem for this approach is the apparent existence of *ser* with adjectives and prepositional phrases:

- (26) a. Juan es feliz/ enfermo /hermoso. Spanish
 Juan SER.3sg happy sick beautiful
 ‘Juan is happy/sick/beautiful (as a person).’
 b. Juan es de Madrid.
 Juan SER.3sg from Madrid
 ‘Juan is from Madrid.’

Roy therefore shows that these apparent adjectives and prepositional phrases are in fact nominals underlyingly in Spanish. Spanish contains ‘nominalized adjectives’, adjectives that can be used as nouns without any overt morphological marking. For example:

- (27) francés; comunista; liberal; militar; científico; viejo; ciego;
 French communist liberal military/soldier scientific/scientist old/elderly blind.(man)
 joven; etc. Spanish
 young/youth

We can talk about, for example *un viejo* ‘an elderly person’, *el ciego* ‘the blind man’, *los jóvenes* ‘the youths’, *tres científicos* ‘three scientists’, etc.

Other ‘regular’ adjectives are not homophonous with nouns.

- (28) importante; mojado; contento; fácil; sorprendido; lleno; recto; alta;. . . Spanish
 important wet glad easy surprised full straight tall

But these adjectives too can have a nominal use, something which certainly distinguishes

them from English where e.g. **He is an important* is unacceptable:

- (29) una importante; los mojados; el sorprendido; unos llenos Spanish
 an important.FEM the wet.PL the surprised some full.PL
 ‘an important one’; ‘the wet ones’; ‘the surprised one’; ‘some full ones’

Whilst there is an internal structural difference between the two types of adjective (Roy, 2013, 153–62), the point about them is that they can both ultimately be analysed as nominals, and as such they can appear in post-copular position with *ser*. That the first type of adjective is a nominal is evidenced firstly by the fact that they allow article insertion, and insertion of the indefinite article triggers the same semantic effects as with ‘regular’ nouns. Secondly, they accept adjectival modifications, despite Spanish generally not allowing adjective stacking (the modification of an adjective by another adjective in the absence of an overt noun). Roy’s analysis of the way that the ‘regular’ adjectives with *ser* are in fact attributive adjectives modifying a null *pro* head also extends to PPs, which are also shown to be modifying a null *pro* head when they are predicates with *ser*.⁷ So we can assume that in Spanish, words that are homophonous with adjectives (and prepositional phrases) can in fact be analysed as nominals.

Roy’s analysis, in line with the general view of this thesis, places the burden of explanation of the interpretive differences in Spanish copular sentences on the flankers themselves, specifically on the predicate, which is either a nominal (including adjectives and prepositions analysed as nominal) or a non-nominal. The copula is held uniformly semantically vacuous. The question—which is really the interesting question for this chapter—then becomes: why do we need two different copulas in Spanish? As Roy (2013, 164) puts it, ‘Why would a language have two morphologically distinct forms of a semantically null item which operate the same function (i.e., that of bearing tense features)?’.

Roy’s answer follows directly from her theory; she simply says that the different copulas align with the different categorical properties of their predicates. Where the predicate is nominal (interpreted as non-dense or maximal) Spanish uses *ser*, and where it is non-nominal (interpreted as situation-descriptive) Spanish uses *estar*. Roy resorts to a morphological model which allows her to claim that *ser* and *estar* are simply allomorphs of a unique copula whose spell-out depends on the syntactic properties of its predicate. Minimal pairs like

⁷*pro* is simply a null pronominal element that UG has long taken to be part of the underlying structure of languages, especially those like Italian that allow null subjects and objects (Rizzi, 1982; Rizzi, 1986) i.e. *pro non vuole cantare* ‘She doesn’t want to sing’. Assuming the existence of such null elements has proved to have explanatory value and contributes to the cross-linguistic similarity of languages underlying their apparent surface differences.

- (30) a. Juan es feliz. Spanish
 Juan SER.3sg happy
 ‘Juan is happy.’ (by nature, a happy person)
- b. Juan está feliz.
 Juan ESTAR.3sg happy
 ‘Juan is happy,’ (at the moment)

would be ambiguous with the copula differentiating the category of the predicate. The point about examples like this is that, as we’ve noted, adjectives in Spanish are homophonous between being ‘real’ adjectives in predicate position or being nominals or parts of nominal phrases. In the case of *feliz*, Roy analyses it as an attributive adjective that modifies a null *pro* in the case of (30a). This gives rise to a non-dense/characterising interpretation. In (30b) *feliz* is simply an adjective, and the sentence is a dense predication, interpreted as a situation description. But the very nature of Spanish adjectives (and prepositional phrases) whereby they have this kind of superficial ambiguity (or perhaps the ambiguity is best described in terms of the kinds of nominals Spanish permits) means that the only sign of the underlying structure of the predicate in these sentences is the copula. So this is what I mean when I conclude that the copula can be a ‘sign of structure’. I mean that in Spanish, it is telling the listener whether the predicate is a nominal or not, which is important in cases where the predicate is an adjective or a prepositional phrase. The copula is what allows the listener to understand whether dense interpretation is intended, or whether a non-dense (or maximal if the indefinite article is present) interpretation is intended.

An alternative to Roy’s view that has been popular in the literature is that *estar* is different from *ser* in having an aspectual content, specifically that it is a perfective verb that takes predicates that are temporally limited and have an endpoint, a view that is intended to capture the contrast between permanent vs. temporary properties, which is traditionally seen as the right description of *ser* vs. *estar*.⁸ A problem with this view (beyond the problems already noted, whereby the temporary vs. permanent distinction just doesn’t line up with a difference in the copula in Spanish, as a matter of empirical fact) is that predicates with *estar* are not compatible with adverbials which are usually allowed with perfective complements, like *recién* ‘recently’, *completamente* ‘completely’, etc. But in any case, even if the proponents of the view that *estar* has aspectual content are right, then this stands as straightforwardly consistent with the view I am promoting here, that the copula can be a sign of structure; in this case the ‘structure’ would be the rather more obviously semantically contentful structure of an aspect head. But I think Roy’s analysis has greater

⁸This is cashed out in terms of *estar* being equivalent to *ser* plus a functional element (which has a prepositional nature in different languages), rather than a lexical property (a feature) by Gallego and Uriagereka (2016).

explanatory power, and stands as a nicer example of the way that the copula is a sign of structure, by signalling the categorical status of the predicate.

5.4 Conclusion

Ever since Aristotle in *De Interpretatione*, the copula has been taken to be ‘a meaningless provider of a tense specification’ (den Dikken, 2006a, 267) or a ‘syntactic hitching post’ for verbal inflectional categories (Pustet, 2003, 2). Den Dikken (2006a, 144) himself says, ‘... it is a meaningless element whose presence in the structure is forced by (morpho)syntactic constraints’. These views are not inconsistent with the two case studies provided here. However, these case studies allow us to abstract even from the tense/verbal inflection-hosting property of the copula to say that it is a sign of underlying structure even in the absence of finite tense, as was the case for the examples I tried to explain in 5.2. The absence of the copula in *A student seems sick* could be taken to indicate that ‘small clause restructuring’ has taken place, thereby disallowing a narrow scope reading for *a student*. (Alternatively, we could say that the presence of the copula indicates that no such operation has taken place.) In fact, in den Dikken’s (2008) theory, the infinitival marker *to* is a tense head, and so indicates the presence of a T in the underlying structure (which is absent in the small clause case), rather as it does in finite copular sentences (as Aristotle noticed). The essence of den Dikken’s (2008) account of the scope differences between *Someone considers every congressman a fool* and *Someone considers every congressman to be a fool* was, following Stowell and Hornstein, that the small clause case creates an environment in which quantifier phrases are unable to escape, so as to allow possible readings where they take wide scope over any quantifier phrases in the matrix clause. In Isabelle Roy’s theory of the *ser/estar* distinction, we clearly saw the way that a language can have a multiple copula system, where the different copulas correspond to spell out reflexes of the different categorical natures of the predicates. When the predicate is nominal, Spanish uses *ser* and where it is non-nominal, it uses *estar*. In both cases, then, the copula seems to simply serve as a sign of the underlying syntactic structure of the sentence. Later, in chapter 7, I’ll question whether it is right to consider both these cases as ‘purely syntactic’ structure. After all, the former case (the embedded small clause vs. infinitival clause in English) relates more obviously to the way that positions within the structure of the sentence are available to move into or not, whereas the Spanish case is more about the categorical nature of the predicates in the sentences. In the latter case, the copula does arguably play a more semantically contentful role, in so far as the categories it spells out can be viewed as semantically contentful in some way. In chapter 7 I argue these categories—i.e. the differences between NumP, ClP, and the non-nominal

categories—relate to the *way we refer*, which should be taken to be a *use* of language more so than a property of syntax. In order to understand what I mean by this, we need to explore the distinction between language as an internal biological property of humans and the use of language, something I endeavour to do in the next chapter.

Chapter 6

Functions

6.1 What am I trying to explain?

So far we have sought to discount a reduction of the copular taxonomy to the two logical representations of predication and identity. We have been assuming that the taxonomy is a categorisation of sentence meanings, but logical predication is not a meaningful notion. Philosophical accounts of ‘pure’ identity and identity statements bear on the meanings of identity statements, though metaphysical accounts of identity only lead to an uninformative notion of numerical identity. The accounts of identity statements we encountered in 3.3 were promising, but their application is restricted to equatives, not specificationals or identificationals. Chapters 4 and 5 served to show that accounts that locate the logical ambiguity between predication and identity in the copula are further hampered by the fact that the copula is not the sort of thing that is meaningful in the requisite way: copulas are not lexical items that can express lexical meanings; therefore, if the meaning of identity in identity sentences is such a meaning, then the copula cannot express it. So in fact any theory that seeks to explain the taxonomy based on different lexical meanings of the copula is fatally flawed according to this picture.

Where do we go now? In this chapter, I wish to continue to pursue the view that the copular taxonomy is a taxonomy of meaning distinctions. To do this, I am going to go back to the source of the taxonomy (Higgins, 1979) and explicate what Higgins’s term ‘function’ might mean. This will lead to my own approach to the taxonomy in the following two chapters. At the very beginning of the introduction to this thesis, I said that the taxonomy of copular sentences cannot be explained moreso than simply described in a better or worse manner. This is the point where I’ll say what I mean by this. I’ll start with Higgins’s own ‘conceptual difficulties’ about the notion of a function 6.2. In order to then understand the notion of a function of a sentence better, we need to get a bit clearer about what language

and communication are in the first place. In 6.3 I outline Chomsky’s notion of I-language and look at what Chomsky had to say about the difference between studying it in its own right and studying its *use*, one example of which is communication. Also in that section I introduce Thom Scott-Phillips’s (2015b) theory, in which ‘communication’ is understood rather differently. Ultimately, however, I don’t think there is any very substantial inconsistency between Chomsky and Scott-Phillips. In 6.4, I show how Scott-Phillips’s account leads to a well-defined notion of ‘function of human communication’ which we can use to derive a notion of ‘function of a sentence’, albeit it is a very general notion that can hardly be used to predict the possible functions of the copular sentence taxonomy. Ultimately, the discussion of Chomsky and Scott-Phillips’s view will lead us to realise that *explanation* is not possible within the realm of the copular sentence taxonomy; we can only hope to describe it. Furthermore, 6.5 also discusses the relationship between ‘sentence function’ and ‘flanker function’, an area where I admit to certain conceptual difficulties. One thing that can be established, though, is that sentence functions cannot simply be deflated into—and thus fully determined by—flanker functions. The final section 6.5.1 shows that the ‘characterisation’ function in Roy (2013) can be fulfilled by different flanker functions in different languages.

6.2 Functions in Higgins

The thesis started with Higgins’s taxonomy, so let’s go back to the beginning and try to understand more deeply what Higgins considered it to be a taxonomy of. For Higgins, the core notions for distinguishing both copular sentences and copular flankers (as in Tables 1.1 and 1.2) were types of ‘function’. But Higgins is upfront about his use of the term ‘function’:

The approach that I have taken here was forced on me by certain conceptual difficulties that I have not yet been able to see my way through. In short, in certain areas of difficulty I am not quite clear about what the real problem is, and hence am unable to devise a means of tackling it. The most direct symptom of this will be when I talk of “functions” of copular sentences, rather than of “meanings” or “readings”. (Higgins, 1979, 205)

He illustrates these conceptual difficulties with an example:

- (1) Which one is Jack Jones?

There are at least two ways this sentence can be interpreted (or ‘used’ as Higgins also puts it). In one, the speaker knows Jack Jones but for some reason cannot see him in a line-up (e.g. due to short-sightedness) and is asking for someone to point him out. In the second

interpretation, the speaker does not know Jack Jones—perhaps he is reading the name off a list of people he needs to talk to—and is asking a hearer to point out the person with this name. Can (1) be considered ambiguous? ‘If it is not ambiguous, how does one account for these different uses of [(1)]? Or are these not even different uses? Is the whole matter linguistically irrelevant?’ (206). For his own part, Higgins is undecided about how to treat such matters, and so resorts to speaking in terms of ‘functional’ distinctions, whilst admitting (as we just saw) that it is a symptom of the limits of his understanding. He draws an analogy with the way that e.g. (2), whilst having the syntactic form of a question, can also be used as an imperative:

(2) Why don’t you shut the window?

(2) therefore has two potential ‘functions’: interrogative and imperative. But it isn’t clear to Higgins in what way a copular sentence like (1) would be ambiguous. He says (207): ‘The factor which seems to be important in copular sentences is more often a distinction between what is known and is familiar and what is not known or is unfamiliar. Because of this, the copular sentence plays an essential role in the communication of new information about known things, but whether this kind of function is to be included in the same paradigm as such functions as questioning and ordering I do not know at the moment’.

As we saw in chapter 1, there are two types of function in Higgins’s taxonomy, the functions of the sentences and the functions of the two flankers. Higgins isn’t systematically explicit about defining these functions, so I’ll just take the cases where he is, to start with.

For the specificational sentence function, Higgins talks about it ‘saying what fulfills a certain condition’ and ‘[t]he Specificational reading in a sense merely says what one is talking about: the Subject in some way delimits a domain and the Specificational predicate identifies a particular member of that domain’ (213). Higgins uses an analogy with lists to explicate the specificational meaning: ‘The heading of a list does not refer to any item at all, nor does the set of items in the list itself say anything about the heading of the list, or indeed about anything. The whole notion of being ‘about’ something is alien to a list’ (Higgins, 1979, 213–4).

With respect to the function of identificational, Higgins says (237) that they are typically used ‘for teaching the names of people or of things’.

As far as I can tell, no such explicit definitions of the predicational or identity sentence functions are given in Higgins. Higgins does spend quite a bit of time explicating the flanker function of the subject of a predicational, the ‘referential’ function, so I’ll now turn to the definitions of the flanker functions. Generally it seems that for Higgins the sentence functions are essentially direct results of the flanker functions:

If, for instance, one takes a constituent that can be Referential as Subject and one that can be Identificational as predicate complement, then the sentence has an Identificational reading. (Higgins, 1979, 264)

So I assume the definitions of sentence functions that Higgins gives (for specificational and identificational that we just saw) are an informal description of what is ultimately reducible to a statement in terms of flanker functions. One such function is the ‘referential’ function. Higgins’s definition of the referential function is essentially that of Wiggins and Geach which I explicated in 3.3.5.1, and so I shan’t say any more about it here. Something else that Higgins stresses (not mentioned in 3.3.5.1) about the referential function is the way in which a sentence is **about** a referential flanker. He says—about the sentence *That man over there is Joe Bloggs*—that ‘it does not seem to me that the name is used Referentially in such sentences—nothing is said about Joe Bloggs’ (Higgins, 1979, 263). This view can also be deduced from what Higgins says about specificational sentences, though his view that the post-copula flanker is not referential is certainly controversial i.e. in a specificational like

(3) What I don’t like about John is his tie.

Higgins denies that *his tie* is referential. This would seem to be at odds with the Wiggins/Geach view of reference: it would not be ‘nonsense’ to ask what *his tie* stood for. Higgins might argue that the second criterion we identified in the Wiggins/Geach view (the need to base a verdict of the truth of the sentence on the identity of the referential expressions) cannot be fulfilled for (3) because Higgins (again, rather idiosyncratically) denies that specificational sentences can be given truth values. The following quote shows that Higgins considers aboutness a criterion for an expression to be referential and that specificational sentences (qua sentences with a list function as said just above) don’t have truth values:

[I]t is not obvious that the notion of having a truth value... is pertinent to lists, for one tends to classify lists as correct or incorrect, complete or incomplete, useful or useless, but hardly as true or false. Just as a list is neither ‘about’ the heading of the list nor ‘about’ the items of the list, so, I would maintain, a Specificational sentence is neither about the Subject nor the Predicate, and therefore neither Subject nor predicate complement is Referential’. (214)

As I say, the idea that neither of the flankers of a specificational sentence is referential is controversial (in fact, I don’t know of any analyses which make that assumption). As is the idea that specificational sentences are not true or false. But the point here is simply that ‘aboutness’, as well as the Wiggins/Geach criteria for referentiality, is part of the definition of the referential function for Higgins.

For the ‘superscriptional’ function (the subject flanker of a specificational sentence), Higgins says: ‘This is the reading that corresponds to the heading of a list’ (Higgins, 1979, 219), which is consistent with his definition of specificational sentences as lists.

Otherwise, I haven’t found explicit characterisations of Higgins’s functions i.e. predicational and identity sentences, and identificational, predicational, and specificational flanker functions. Extrapolating from what Higgins does explicitly say, we could perhaps deduce that the predicational flanker function is the inverse of the referential one. So this would mean that for predicational flankers: (1) It would be nonsense to ask which individuals they stand for. (2) We would not need to know which individuals they stand for in order to judge the truth of the sentences they are in (the Wiggins/Geach criteria, inverted). (3) Sentences containing predicational flankers are not **about** those flankers.

For the flankers, that leaves the ‘identificational’ function of the predicate of ‘identificational’ sentences and the ‘specificational’ function of the predicate of ‘specificational’ sentences. Again, extrapolating from Higgins, we can follow what he says about the superscriptional function. If the superscriptional function corresponds to the heading of a list, then the specificational function would correspond to an item on a list. For the identificational function, we could say that it would correspond to the name of an object or a person, given the function of identificational sentences as ‘teaching the names of objects or people’.

That just leaves definitions of the predicational and identity sentences. Given the way readings seem to follow from flanker combinations for Higgins (see the quote just above), we can say that a predicational sentence is one where something predicational (i.e. which cannot be the answer to a ‘which individual is it?’ question) is said about something referential (i.e. which can be the answer to a ‘which individual is it?’ question, the knowledge of which can determine the truth of the sentence).¹ An identity sentence is a sentence which is about both its flankers, which are both identifiable in advance of the truth of the sentence. This last definition seems the least satisfying to me as a definition of identity, as other verbal sentences contain referential arguments whilst not being identities (though perhaps the fact that one is embedded in a predicate means that it is not referential):

(4) Mary kissed John.

So this, I think, gives an idea of what a function is for Higgins. In many ways I agree with Higgins’s approach and follow it in my own analysis. So in what follows, I wish to try and

¹What about ‘ascribing a property to an object’ as a function of predicational copular sentences? Higgins does not deal in these terms, as far as I’m aware. If it is understood as the functional correlate of syntactic predication, then it applies universally to all copular sentences, and is essentially vacuous for the purposes of discriminating copular sentences (whether it would serve any other purpose is debatable). As a statement of metaphysics, it may have more merit, see chapter 7.

Table 6.1: Functions of copular sentences and flankers, derived from Higgins (1979)

<i>Sentence functions</i> (= subject flanker function + predicate flanker function)	
Identificational	Teaching the names of people or things
Identity	A sentence about both its flankers, which are both identifiable in advance of knowing the truth of the sentence
Predicational	Referential flanker + predicational flanker (? = Something predicational said about something referential)
Specificational	Saying what fulfils a certain condition / Identifying a member of a delimited domain / Listing
<i>Flanker functions</i>	
Referential	(1) Not ‘nonsense’ to ask what individual a referential flanker R stands for, (2) Truth of the sentence is based on identity of R, (3) Sentence is about R
Superscriptional	The heading of a list
Predicational	(1) ‘Nonsense’ to ask what individual a predicational flanker P stands for, (2) Truth of the sentence not based on identity of P, (3) Sentence is not about P
Specificational	An item on a list
Identificational	The name of an object or person

do what Higgins did not attempt, which is to justify talking in this way about sentence and flanker functions. Higgins’s copular taxonomy is, I believe, at least partly a taxonomy of functions. This discussion can be given in table form as in Table 6.1.

6.3 Prerequisites to giving the function of a sentence or parts of a sentence

6.3.1 What gets used?

So we want to know what the function of a sentence or parts of a sentence is. A function is a use of something. In order to say that a function is a use of language, we need to try and be as clear as possible about what ‘language’ is, such that it can be used.

For Chomsky, sentences are a consequence of an evolved Universal Grammar (UG) that is universal in the species and which is the initial state that—given exposure to a particular language like English, French, or Swahili—then develops into an individual’s I-language over the course of first language acquisition (Smith, 1999, 31–4).² An I-language is essentially the

²The term ‘I-language’ first appears in Chomsky (1986); prior to this—in works like Chomsky (1980)—he

initial UG state with certain modifications. This I-language is then the basis for generating sentences of English, French, Swahili etc. (where English, French, Swahili are E(xternal)-languages). I-languages are unique to an individual, though there will be considerable overlap between individuals whose I-languages generate sentences of a given E-language, in such a way that they can understand one another.

What are examples of specific components of UG and an I-language? UG contains—at least—the most basic properties of I-language that enter into all I-languages: operations like Merge, Move (which may, as indicated in section 2.3.7, simply be ‘internal merge’, a kind of Merge that ‘comes for free’ given Merge/external Merge, (Chomsky, 2004)), and Agree (the operation whereby certain interpretable semantic features of number, gender, person, and case (if case can be called a semantic feature) on some lexical items agree with uninterpretable such features on other lexical items in a given derivation, driving movements/internal merges).

There has been more controversy about how to specify I-languages. Until recently, the ‘Principles-and-Parameters’ (P&P) theory of Universal Grammar had it that that certain principles and parameters were part of Universal Grammar, and that it was then the setting of the parameters upon exposure to environmental linguistic stimuli that generated the final state I-language in an individual (Chomsky, 1983). So, for example, a typical P&P work like Baker (2001) lists a number of such principles and parameters. An example of a parameter would be the ‘head directionality’ parameter, which essentially sets an individual’s I-language such that heads either precede or come after their complements. This then accounts for—amongst other facts—the fact that languages are all either VO or OV i.e. verbs either precede objects or come after objects. Nowadays, the Minimalist Program has done away with such parameters. Particular theories within the general scope of the program tend to situate language variation in terms of varying featural specifications of lexical items. It is features that drive movements, so varying featural specifications will give rise to different word order effects. Kayne’s (1994) theory of how syntactic structures are linearized is a Minimalist-style theory that supersedes parameters like the ‘head directionality’ parameter; it claims that all underlying syntactic structures are head-initial, but that varying syntactic movements in different languages (on the basis of varying featural specifications of lexical items, for example) give rise to the surface-level differences whereby heads precede their complements in some languages and come after them in others. The only other thing to say about it is that I-language interfaces with two systems, the ‘conceptual-intentional’ (i.e. meaning) system and the ‘sensory-motor’ (i.e. pronouncing, signing etc.) system.

So for the theory of Universal Grammar, ‘using language’ means using I-language, using used ‘the ambiguous term’ (Smith, 1999, 216) ‘grammar’.

this internal, innate (albeit modified upon environmental linguistic exposure), biological endowment. Summarising Chomsky's views with respect to the commonsense notion of a 'word' or 'sentence', McGilvray (Chomsky & McGilvray, 2012, 160) says that they can 'refer to... things and circumstances in the world'. As such, they are 'about' things and circumstances, but McGilvray is at pains to point out that words and sentences are not about things 'all by themselves' but only once they are used 'in the relevant ways (successfully, on some accounts)' and, 'in this sense, reference and 'aboutness' come about because that is how we happen to use some words and sentences, sometimes'. This then leads to a statement of the overall picture that the last few paragraphs gives rise to:

[A]n intensional or theoretical specification of an I-language may be a construct in the mind of a linguist, but for Chomsky, it is also a description of a 'real' state of an 'organ' in a human mind. That state is a developed state of UG, developed in accord with biophysical constraints on a possible language. An I-language so described is assumed to be 'the real thing,' the proper object of linguistics thought of as a natural science. The sentences produced by a person, necessarily with the aid of whatever 'performance' systems the language faculty cooperates with, is an epiphenomenon, and only that. (Chomsky & McGilvray, 2012, 168)

The notions of reference, aboutness, words, and sentences, then, are all uses of I-language for Chomsky.³ This is a crucial point to note. It is often overlooked how deep use goes for Chomsky. The focus tends to be on Chomsky's assertions that communication is not *the* function of language, merely one function⁴:

³Hinzen (2006, 175) ostensibly agrees with this, saying 'sentence' is 'no more than a descriptive artefact: a sentence is simply the extended projection of the verb that heads the sentence, an IP, and appears to share much of its internal construction principles with the nominal domain'. However, I'm not sure Hinzen is right to say that 'sentence' becomes a 'descriptive artefact' under this analysis. It rather becomes synonymous with IP, which I believe is mooted by Universal Grammar as a property of the natural world; it enters into the description of an individual's I-language for a start. 'Copular sentence' is more of a descriptive artefact than 'sentence'; it is certainly not cross-linguistically universal. Whether Chomsky's notion of sentence is the same as Hinzen's is unclear to me.

⁴Chomsky, in fact, is more deeply sceptical of the possibility of explanations of language in terms of natural selection:

The processes by which the human mind achieved its present stage of complexity and its particular form of innate organization are a total mystery...it is perfectly safe to attribute this development to 'natural selection', so long as we realize that there is no substance to this assertion, that it amounts to nothing more than a belief that there is some naturalistic explanation for these phenomena. (Chomsky, 1972, 97)

Subsequent authors, such as Pinker and Bloom (1990), have explicitly disagreed with this stance and sought to forge a synthesis of UG and Darwinian evolutionary theory in which increasingly successful communication is considered to correlate with increased fitness of the organism. But such gradualist accounts of language

It's perfectly true that language is used for communication. But everything you do is used for communication—your hairstyle, your mannerisms, your walk, and so on and so forth. (Chomsky & McGilvray, 2012, 12)

In fact, Chomsky believes that '99.9 percent of its use is internal to the mind'. But how are we to understand this? In the next section, I want to argue that the way *not* to understand it is as saying that communication simply means 'using language'. Chomsky seems to be aware of this; what he says implies that there must be more to articulating communication than simply 'using language'. We'll see that to do this would literally be a category error. This reflects more recent work into the nature of communication systems in biology generally, and in humans in particular. The problem with Chomsky's work is that he nowhere refines what he means by 'communication', so that his claim that communication is a peripheral use of language is impossible to evaluate. In the next section, then, I will outline what human communication is such that I-language could be used to do it, and consequently to refer and produce words and sentences.

6.3.2 I-language can't simply be used *tout court*

For Scott-Phillips (2015b), communication cannot be stated as 'the use of I-language'. Rather, it is the result of an evolved communication system which is quite distinct from I-language. Communication is not a use of I-language; it is the use of a communication system. Language in Chomsky's UG/I-language sense gives rise to a coding capacity which enhances human communication. So rather than saying that I-language is used to communicate, we have to say that a communication system uses I-language, in order to enhance communication. Note that under this view, it is in fact the communication system that—on an adaptationist view of evolution—must be justified in terms of a function. We have to ask what function such a communication system serves. This is different to Chomsky, who simply seems to accept that communication is an ultimate function in its own right. Another way of stating Scott-Phillips's position is—as he puts it himself—'... the common assumption that the linguistic code [read: I-language] makes linguistic communication possible is simply false'. This, then, means that we can't say that the kind of communication that humans engage in can simply be a *use* of an I-language; on Scott-Phillips's account human-style communication is a particular (and biologically unique) kind of thing, and notably different to the only other definition of communication, 'code-model' communication, which is indeed enabled merely by the existence of codes. I-language gives rise to a partic-

evolution run into difficulties from paleo-anthropological evidence which suggests language evolved recently (c.200kya) and suddenly (Tattersall, 2002; Tattersall, 2017).

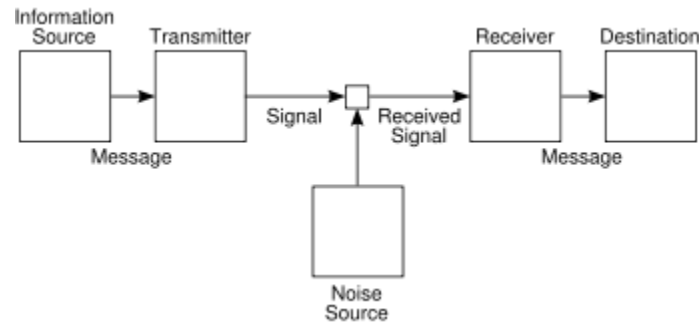


Figure 6.1: The information-theoretic approach, Shannon (1948)

ular kind of code system (which, again, may well be biologically unique; it is a code system with particular properties of hierarchy and structural distance, for example). But to view I-language as enabling communication via an undefined notion of communication in saying ‘using I-language for communication’ is to literally make a category error: I-languages could, on their own, only ever enable code-model communication not human-style communication. Scott-Phillips’s view merits being taken seriously. For one thing, he articulates clearly the nature of communication, a concept which has been well-studied in biology and linguistic pragmatics.

6.3.2.1 The two, and only two, possible communication systems in philosophy and biology

I won’t go into all the details of the theory of communication here, but just the basics, so that the reader can approximately understand the claim in the previous section.

‘Communication’ has been technically defined, notably within biology, but also within the field of information theory (the field information technology is based on). Only *two* kinds of communication have ever been discovered. If there is a philosophical theory of a third kind of communication, it has not so far been proposed (Scott-Phillips, 2015b, 13). The most intuitively obvious kind of communication is ‘code model’ communication.

The diagram in figure 6.1 outlines the code model, as understood in the pioneering work of Shannon (1948). A message is transmitted (i.e. is a signal) from one end of a conduit and received at the other end. Code model communication is, crucially, associative. In the natural world, there is an association of a state of the world with a behaviour, which behaviour then causes an associated response elsewhere either internal to or external to a given organism where the signal is received. In that case the ‘code’ is the behaviour associated with the particular state of the world. To give an example, in bacterial ‘quorum sensing’:

... individual bacterial cells produce small diffusible signal molecules in particular

local environments. . . These molecules bind to cognate receptors on other bacteria, and in doing so effectively reveal the presence of the bacterium in the local environment. When the concentration of these molecules reaches a threshold level (the quorum), it triggers population-wide behaviour, such as gene-regulation. (Scott-Phillips, 2015b, 4)

So a state of the world ('particular local environments') is associated with a behaviour (the production of small diffusible signal molecules) which is the signal. This is received by the receptors of other bacteria which associate a certain threshold of such molecules with a behaviour, such as 'gene-regulation'. The code for such communication can be given in the form of encoding and decoding algorithms:

- (5)
- a. Encoding algorithm: If in environment X, produce molecule Y.
 - b. Decoding algorithm: If concentration of molecule Y is above threshold t, then perform behaviour Z.

Arguably, almost all communication in nature besides human communication is of this sort, though humans also exhibit it. For example, the areolar glands of new mothers produce a chemical that attracts the newborn towards it, in order to feed (Doucet *et al.*, 2009). The code can also be given in the form of algorithms:

- (6)
- a. Encoding algorithm: If just given birth, secrete chemical X.
 - b. Decoding algorithm: If X is detected, move towards it (and suck it!).

Again, we see here a state of the world associated with a behaviour (the signal: secretion of a certain chemical), the detection of which is associated with a behavioural response. This is an example of a human 'chemosignal', of which there are several others. For example, 'women's emotional tears' have recently been studied and shown to contain a chemosignal that has the effect of reducing sexual arousal in men (Gelstein *et al.*, 2011). Again, a state of the world (the state that produces tears in women) is associated with a behaviour (the production of tears) which are a 'chemosignal' that associates with a response in men (reduction of sexual arousal).

So much for the code model, which may well account for almost all the communication that goes on in nature outside human communication. The other kind of communication (and again it must be stressed the *only* other kind of communication that anyone has ever thought of) is termed the 'ostensive-inferential' model of communication by Thom Scott-Phillips (2015b), 'ostensive communication' for short. He introduces this by way of examples: If we're both preparing dinner together, and I push a pile of unchopped vegetables towards you and a knife, I am clearly communicating some kind of message (i.e. 'Chop the

vegetables!'). Now, clearly some communication has taken place, but there is no code in operation here; the behaviour is not *associated* with the intended meaning, i.e. there is nothing in the environment that the production of my behaviour is associated with. I could have produced other behaviours, and other circumstances could produce the same vegetable-pushing behaviour. I might have produced a completely novel behaviour, or at least one that you haven't seen me produce before, whilst still expecting you to understand my meaning.

This kind of communication seems to be working on mental representations (e.g. beliefs, assumptions, goals) and changing them in some way. Scott-Phillips calls this an 'informative intention', which he defines (8) as 'a signaller's intention that the receiver change their representation of the world in response to the signaller's behaviour'. The content of an informative intention is 'the changes that the signaller wants to make to the receiver's mental representations'. To take another simple example Scott-Phillips gives: in a restaurant, a customer catches a waiter's eye, holds their coffee cup up and tilts it slightly. The content here is fairly obvious: something like 'I want a refill'. This is the change in the waiter's representation of the world that the customer wants to effect.

But there needs to be a second kind of intention in order for these examples to work. The customer in the restaurant needs to communicate the very fact that she wishes to communicate at all. After all, coffee cups get tilted all the time for non-communicative reasons. How does the waiter know that this isn't the 'meaning' of seeing a tilted coffee cup? The tilt has to be done in an *ostensive* way: the customer establishes eye contact and does a particular kind of stylised tilt once she's established she has the waiter's attention. Without recognising the 'communicative intention', the waiter would probably never infer the informative intention, and so no communication would occur.

Another example: Peter sees Mary eat berries in a 'somewhat exaggerated, stylized way', patting her tummy as she does so.

- (7) a. Informative intention: Mary thinks the berries are tasty
 b. Communicative intention: Mary wants to communicate this fact to Peter (the communicative intention is always the same i.e. the intention to communicate a particular informative intention.)

Crucially, note that if Mary just ate the berries enthusiastically (reflecting the fact that she was enjoying them), Peter could still infer that the berries are tasty (or at least Mary finds them so), but I think it is clear to see that this would not constitute an act of communication. I can infer from the fact that my outside bin is empty that the binmen have been to collect the rubbish already, but there clearly wasn't any kind of communication involved there, merely a state of the world from which inferences can be made.

Comparing the code model and the ostensive model directly, we can see that in a very general way both are designed to cause changes in the receiver's behaviour, but they do this via different causal mechanisms. The code model works on the basis of associations. The signaller *encodes* some specific content which is designed by evolution to trigger associative responses.⁵ The causal mechanism of ostensive communication is 'metapsychology', i.e. the ability to think about each other's thoughts, reason about each other's reasons etc.⁶ Rather than encoding some specific content, *the signaller provides evidence that she has a communicative intention to express* some specific content. And rather than being designed by evolution to trigger associative responses, ostensive communication is designed by evolution to change mental states.

Now, where does language in a UG sense (or even just a pre-theoretical sense of words and sentences) enter the picture in the ostensive communication model? Evidently, all the examples given so far could be augmented with 'linguistic codes' (to use Scott-Phillips's preferred term) which are 'reliably, typically associated with certain referents'. In the case of the coffee-cup tilting, the customer could say to the waiter 'I'd like some more coffee please'. In the berry-eating case, Mary could say to Peter 'These berries are very tasty'. As Scott-Phillips puts it (17–8): 'In this sense, linguistic communication is a very important special case of ostensive communication, one in which ostensive communication is augmented by the linguistic code. . . I can nonverbally but ostensively point to any of the objects in this room, but with language I can refer to any object in the world. I can make a request of others by pushing unchopped vegetables in their direction, but with language I can make requests about things remote in time and space.'

The inspiration for all this is the work of the philosopher Paul Grice, e.g. Grice (1957; 1989), and another way of presenting the two types of communication outlined above could have been in terms of Grice's distinction between 'natural' and 'non-natural meaning', and the way that natural meaning enters into 'natural codes' in code model communication, as distinct from the 'conventional codes' that allow us to convey non-natural meaning.⁷ The work of Tim Wharton, in particular, has connected Grice's work on meaning with the unique

⁵The solution to the apparent chicken-and-egg problem whereby this arises in nature—the story of 'cues', 'coercion', 'ritualization', and 'sensory manipulation' is an interesting discovery of recent biology in its own right, but not something I'll have space to go into here; see Scott-Phillips (2015b, ch.2).

⁶The cognitive mechanism underpinning ostensive communication is referred to as 'recursive mindreading' by Scott-Phillips and again is itself a fascinating discovery of modern cognitive psychology, but there's no space to go into the details here. It is essentially a recursive theory mind i.e. the ability to believe that you intend that I believe that you intend that I believe something. It is, in fact, precisely this, a representation that requires five levels of embedding an informative intention. At this point, and only at this point, does ostensive-inferential communication occur. See Scott-Phillips (2015b, ch.3).

⁷And the general view that meaning derives from the use of language which Grice builds on to some extent is of course most closely associated with the work of the later Wittgenstein, e.g. Wittgenstein (1953).

kind of communication that humans are capable of (Wharton, 2003; Wharton, 2009). The way Grice (1989, 92) sums up his account of non-natural meaning is:

U meant something by uttering x is true iff, for some audience A, U uttered x intending:

- 1.) A to produce a particular response r
- 2.) A to think (recognize) that U intends (1)
- 3.) A to fulfil (1) on the basis of his fulfilment of (2)

(1) is equivalent to Scott-Phillips's informative intention, (2) to his communicative intention, and (3) is essentially the requirement for some kind of ostension.

'Meaning' in code model communication can be thought of as (in biological terms) an 'ultimate level' explanation of an animal's signalling behaviour, the function of this behaviour. Grice's account of non-natural meaning is a proximate explanation of human behaviour, as it only refers to mechanisms of meta-psychology; there is, however, an ultimate explanation/function of this ostensive communication, which is given below. So it is quite innocent to say that mating calls mean that an animal wants to mate, as this is their biological function, causing in others a willingness to mate. And we can say that a vervet monkey's call can mean 'eagle' as long as we recognise that the monkey isn't trying to change the other monkey's mental representations in any way as a consequence of the signal; rather, it 'meaning eagle' just reflects the naturally selected function of the call which is to elicit the response in the monkey's fellow monkeys of taking action to evade an eagle (by diving into the undergrowth, for example). 'Meaning' in either a code model or ostensive communication model is about 'doing things to others in a designed way' (Scott-Phillips, 2015b, 26): non-human animal signals are designed by natural selection to change behaviour; human signals are designed by human intentions to change mental representations (the ability to do this also, of course, a consequence of natural selection).

6.4 The way in which I-language *can* be used and the function of a sentence

We now have the ideas in place to see how it might make sense to say that I-language can be used for communication, and to talk about the function of a sentence.

I shall assume that Scott-Phillips is basically right. We evolved a capacity for ostensive-inferential communication, which is founded on the cognitive mechanism of a recursive theory of mind. Now, if this is right, then it is necessary to say what the function of *this system*

is. Notice that this is quite different to the way that we previously talked about a function of I-language being communication. We simply can't say that any more. Now we say that we have a communication system, which itself has a function, which uses I-language (but doesn't have to) in order to fulfil its function.

What then is the function of ostensive communication?

6.4.1 A function for ostensive-inferential communication

Firstly, communication is not a single trait, but rather two distinct, though interdependent, traits of signal production and reception, ostension and inference in Scott-Phillips's terms. But the interdependence complicates matters with respect to thinking about what would constitute good design for communication. In terms of signal reception we can say that the listener wants to 'come to believe or know or understand something worthwhile that they did not believe or know or understand beforehand' (Scott-Phillips, 2015b, 138). The speaker also wants the listener to achieve this goal. But, '[t]he rub is that these may not be the same things that the listener cares about'; the things that the speaker may want the listener to know may not be the same as the things the listener may want to know.

Listeners must draw relevant conclusions from stimuli. If they hear the sentence *This is a spoon*, they shouldn't waste cognitive effort drawing the conclusion that the object I am holding is not a knife, fork, potato, lobster or, indeed, an indefinite number of other things. This is a version of the 'frame problem', which originated in artificial intelligence as 'the challenge of representing the effects of action in logic without having to represent explicitly a large number of intuitively obvious non-effects' (Shanahan, 2016).

Speakers too must limit their efforts: 'If listeners filter for relevance, then telling them everything regardless of relevance is a waste, since it will not have any effect' (Scott-Phillips, 2015b, 139). So both listeners and speakers must have cognitive systems that are designed to maximize relevance; for listeners, this is the relevance of the inputs they receive; for speakers, this is the relevance of the stimuli they produce for their listeners i.e. they should produce stimuli if they are relevant for their audience. So to be clear, we have now stated what it would be for ostensive-inferential communication to be well-designed, to be an adaptation: it must maximize relevance for listeners and speakers.

So the big question is whether this is in fact the case, whether human communication does maximize relevance in this way. The answer is yes (Scott-Phillips, 2015b, 57–63). Ostensive communication works in accordance with the principles of relevance theory (Sperber & Wilson, 1995; Wilson & Sperber, 2004), which ultimately re-cast Grice's (1975) maxims of conversation in terms of a principle of optimising the cognitive effort of producing and

processing a signal in relation to the ‘cognitive effects’ i.e. the benefits of doing so. So Scott-Phillips concludes that ‘humans are indeed adapted for ostensive communication, that is, to express and recognize communicative and informative intentions, and in doing so better navigate their social environment’ (Scott-Phillips, 2015b, 139).

This itself must also prove to be useful; it must have been selected for by natural selection. And so we must ask about it: Why would it have been so selected? In Millikan (1984)’s terms, the *direct* function of a trait is what is historically responsible for the reproduction of it. So why did ostensive communication evolve as an adaptation to express and recognize communicative and informative intentions, and in doing so better navigate one’s social environment? The direct function of ostension is the manipulation of others’ mental states. The direct function of inference is the mindreading of others’ mental states: ‘ostensive communication is a form of extended social navigation, in which signallers try to mentally manipulate their audience, and audiences try to mindread their signallers’. Derived sub-functions of this might be the sorts of things that have been suggested to be the function of communication in the literature: gossip, mate attraction, sexual competition, planning for hunting, the advent of throwing, the emergence of politics etc. (Scott-Phillips, 2015b, 140).

6.4.2 Reconciling I-language with the function of ostensive-inferential communication

The final statement of the function of ostensive communication, then, is that it is ‘a form of extended social navigation, in which signallers try to mentally manipulate their audience, and audiences try to mindread their signallers’.

The way in which I-language gets used for communication, then, really amounts to the way that it gets used by the system for ostensive communication. It interfaces with this system in rather the same way as it interfaces with the conceptual-intentional system and sensory-motor system (see above, section 6.3.1).⁸ This, incidentally, is also a way in which language has been said to be *used*; Hinzen has called the use of I-language by the interfaces an ‘internalist functionalism’ (Hinzen, 2012a), see also Hinzen (2006, 216–7, n.31).

Also, it is unclear whether we ever use I-language *not* in the context of ostensive communication.⁹ What would be an instance of this? There are certainly conceptual difficulties

⁸What the precise nature of the interaction of all these systems is (I-language/syntax, the C-I/S-M systems, and ostensive communication) is unclear to me.

⁹It’s interesting to note that Scott-Phillips would probably agree with Chomsky’s (Chomsky & McGilvray, 2012, 12) assertion that ‘everything you do is used for communication—your hairstyle, your mannerisms, your walk, and so on and so forth’ (which Chomsky frequently brings in as an argument to say that communication is not the, or even a, main function of language). But Scott-Phillips’s theory allows us to see *why* this is the

here, in understanding what an unused I-language is exactly, and whether it is ever not filtered through the system for manipulating and recognising informative intentions. These questions have not been broached yet, as far as I am aware. For now, I shall assume that I-language on its own is something like a biological capacity for hierarchically organising perceptual stimuli. I shall follow Chomsky in considering that reference and sentences are uses of I-language. But I shall follow Scott-Phillips in considering them to be uses of I-language by ostensive communication, and as such instances of the overall function of ostensive communication, viz. to mentally manipulate one's audience as a form of extended social navigation. Now, whilst I think there is some substance to this analysis, it is of course very general, and with respect to the taxonomy of copular sentences, they could all be classed as mental manipulations of an audience for extended social navigation.

Note that manipulating mental representations is a more general notion of communication than what Chomsky sometimes seems to have in mind. Chomsky thinks that a 'substantial' notion of communication involves 'conveying information':¹⁰

So if you're at a party there's a lot of talk going on. But the amount of communication that's going on is minuscule; people are just having fun, or talking to their friends, or whatever. (Chomsky & McGilvray, 2012, 13)

This is consistent with the sociological literature, which has analysed many aspects of conversation, such as small talk, in terms of achieving social purposes like bonding or managing personal distance. For example, Thornbury and Slade (2006, 25) say: 'Conversation is the kind of speech that happens informally, symmetrically and for the purposes of establishing and maintaining social ties'.

In this line of thinking McGilvray claims the following as linguistic functions:

Words and sentences... are used to classify, describe, refer, insist, cajole, make claims, and—a favorite of many, and often claimed to be essential in some way—communicate. (Chomsky & McGilvray, 2012)

case: any action that can be performed ostensively can be used for ostensive communication, and it isn't obvious that there are actions under conscious control that cannot be performed ostensively.

¹⁰Incidentally, biology seems to be clear that information transfer cannot be a definition of communication. This is because it cannot distinguish between 'cues' and 'communication'. 'Cues' are a stage on the way to the development of communication in which information that is incidentally produced by one animal is used by another. For example, the appearance of a predator on the horizon can be a cue to another animal to behave in a particular way. Dogs' territory-marking with their urine is taken by Scott-Phillips (2015b, 34–5) to be an instance of a cue-first evolution of communication. Originally dogs involuntarily urinated as a fear response, which information was used by other dogs. Over time, they have evolved to control urine production, and it is only at this point that we can talk about communication, and say that urinating can be truly a signal for dogs. Though it is important to note that information transfer was occurring even in the pre-communicative stage. So information transfer is not definitional of communication, as Chomsky seems to think.

Later McGilvray adds to his list the functions ‘to describe, speculate, convince, classify, question, communicate, and so forth...to refer and assert, complain and praise, and— anomalously, assuming that words come ‘out loud’—silently think and ruminate’. But notice that these all (apart, perhaps, from silently thinking and ruminating) seem like perfectly good cases of manipulating someone’s mental representations as a form of extended social navigation.

6.5 We can only describe, not explain, the taxonomy, and some ‘conceptual difficulties’

Ultimately such functions are uses of language, they are actions, and, as McGilvray notes (Chomsky & McGilvray, 2012, 159), ‘action is free’, which is to say (I believe) that it is governed by free will. As such, why any given action occurs may not be explicable in terms of a deterministic scientific theory. Whether there’s any hope for developing a theory of the range of *possible* actions in the domain of linguistic functions is unclear to me, but I’m certainly not aware of any such theory. Theories of ‘functional grammar’ e.g. (Dik, 1989; Halliday & Matthiessen, 2004) do not, from what I can tell, predict possible functions of sentences more so than work bottom-up from sentences to a characterisation of functions (rather as Higgins did). About this, McGilvray says:

There is nothing wrong with describing use, of course. The problem is, as Wittgenstein pointed out long ago, you cannot find in these highly context-sensitive and variable descriptions of the ways in which people use language to serve all sorts of purposes the regularities that any serious form of theorizing requires. Lewis [in Lewis (1975)] and others needed and need to be disabused of the illusion of uniformity in the use of natural languages and told that if they want to construct theories of language at all, they must look to language not in use, but to languages as natural objects that allow for use. (Chomsky & McGilvray, 2012, 165)

So whilst this isn’t something that can produce a scientific theory as such (unlike the theory of I-language), it isn’t perhaps a completely fruitless enterprise, and it would perhaps be unfair to call it a completely unscientific project.¹¹ It has some taxonomic value at the

¹¹But I accept it is not pursuing ‘the fundamentally Galilean concept that you ought to try to discover ideal cases and see if you can develop principles that’ll then ramify to account for the world’ (Chomsky, 2009), something which is not possible in the realm of copular sentences, which are fundamentally a part of human action rather than a part of the biological world about which scientific theories can be formed.

very least. As McGilvray says, the exercise may offer some data, ‘... data that the scientist of language and mind can and probably should take into account’ (Chomsky & McGilvray, 2012, 164).

So this is the place to be totally explicit about the way that this thesis can offer no more than a *description* of uses of I-language (by ostensive communication). This isn’t to denigrate the status of description; on the contrary, the claim here is that this is all that is appropriate. As such, then, I won’t claim that it is a scientific theory of copular sentences (in line with Chomsky, I don’t think such a thing is possible). But I believe this is the case for anyone that wishes to say something about Higgins’s taxonomy. A syntactic theory of copular sentences is possible, I believe, about which I’ll say a few words in chapter 7. But if we’re interested in the meanings of copular sentences (beyond the extensions of the entities referred to, which may well be explicated by a metaphysical theory, see 7.3), then description is the best we can aim for.

As I said earlier, the notion of a sentence as an instance of mentally manipulating one’s audience as a form of extended social navigation doesn’t take us very far in helping to characterise the range of copular sentence functions (bear in mind that we don’t even need to produce sentences to achieve this; we can just push unchopped vegetables at each other). At this point I’ll admit to another conceptual difficulty. It isn’t clear to me how to determine what ‘the one function’ would be for a sentence, and so I’m not going to try to do this. Rather, I’m going to somewhat follow in Higgins’s footsteps and talk about the functional properties of the flankers of copular sentences and then assume that there is some kind of relationship between these properties and ‘the function of the sentence’, if there is one. Remember that Chomsky said that referring is a use of language. This is crucial. It also tallies with what Scott-Phillips says about the ‘linguistic code’, which I am equating with the raw material products of I-language here: linguistic codes are ‘sounds and/or gestures that are reliably, typically associated with certain referents’ (Scott-Phillips, 2015b, 17). Scott-Phillips is also explicit elsewhere that conventions such as the reliable association between a sound/gesture and a referent should indeed be analysed at the level of component parts of sentences: ‘[I]t’s not sentences that are conventions, it’s the component parts, including the structural elements. This includes words, phonological patterns, morphosyntactic operations, and so on. Critically, all of these are tokened. It’s in this way that languages are sets of conventions’ (Scott-Phillips, 2015a).¹² This, I think, helps to justify giving priority to flanker functions

¹²Another conceptual difficulty: it isn’t clear to me that Scott-Phillips’s ‘informative intentions’, if they can be generalised, would be distinct from what most people would think of as the function of a sentence. An informative intention, recall, is a mental representation (or intention to change one) e.g. in the case of the coffee cup tilt, the informative intention can be stated as ‘I want a refill’. ‘I want a refill’ is a token, but if we generalised it to something like a ‘request to do something for me’, we get something that looks like a

over sentence functions. The relationship between functions and traits is quite complex in biology quite generally. Take the blood-pumping function of the heart. This would qualify as the function of the heart if we understand function in Millikan (1984)'s sense of the function that is historically responsible for the heart's reproduction. But what is the relationship between blood pumping and the traits that enable the function (muscle contraction, cavities filled with blood, connections to the circulatory system)? It isn't clear to me here, and I certainly wouldn't know about the relationship between the function of a sentence (to the extent that one can be given) and the component properties of the sentence; this extension from a non-intelligent to an intelligent system may not be legitimate at all.

As I said, I will follow Higgins who assumed that his sentence functions were synonymous with a statement of the flankers ('If, for instance, one takes a constituent that can be Referential as Subject and one that can be Identificational as predicate complement, then the sentence has an Identificational reading' (Higgins, 1979, 264)). In any case, there sometimes appears to be a very transparent relationship between the purported function of the sentence and that of the flankers. The 'listing' function of a specificational appears synonymous with a description of the flanker functions i.e. a list heading with a list item. In what I say later, there will also be an apparently quite transparent relationship between the flanker description and any putative sentence function. Even though I don't know how to determine the relationship between a sentence function and a flanker function, in the next section we'll address the question of whether it is a deterministic relationship, i.e. whether the flankers uniquely determine the sentence function or not.

6.5.1 The relation between sentence and flanker function: two possible ways to characterise?

There is something that can be concretely said about the relationship between a sentence function (if there is such a thing) and its sub-parts. If we think that it would make sense to consider the flanker properties of a copular sentence as combining to produce a sentence function, then there are at least two ways to understand this: (1) The sentence function can be deflated into—and is thus fully determined by—the flanker properties (2) there is a a looser relationship; one function could be realised by sentences with different flanker properties, but perhaps there is some restriction on what properties can be used to realise a given function.

We'll now briefly look at a case study from Irish Gaelic that eliminates option (1). It

function of some kind. Having said that, the goal of a goal-directed action seems to always be describable in these terms i.e. in more or less specific terms. I could be taking the top off my bottle of coke, taking the top off a bottle, removing some sub-part of an object etc.

will not be possible to say that a function can be fully determined by its flankers. My own conclusion about this is that we should give up trying to assign unique functions to sentences, rather than opting for option (2), though I will suggest at the end that a ‘weaker’ notion of ‘inter-translatability’ may at least be worth exploring.

In chapter 1, we encountered Isabelle Roy’s extension to Higgins’s taxonomy, sub-dividing the predicational class into three sub-categories: characterisation, definition, and situation-description. These appear to be ‘functional’ labels of some kind (Roy doesn’t expand on these terms anywhere that I can find), but they are clearly to be understood as reducible to the semantic descriptions she gives which are then correlated with syntactic descriptions (such that Roy could plausibly claim that she has a syntactic *explanation* for the three types of predicational sentence she identifies). So, it is rather like what Higgins does—sentence functions are a function of flanker properties—except that for Roy it seems fairer to say that, in her terms, she attempts to reduce the flanker properties to syntactic descriptions, rather than other ‘functional’ properties, as Higgins does (but see just below for a tension between Roy’s work and Chomsky’s view of reference that we encountered earlier in 6.3).

Let’s focus on her ‘characterisation’ class. To characterise is to produce a sentence like B’s:

- (8) A: Que fait Jean dans la vie?
 ‘What does John do?’
 B: Il est prof.
 ‘He’s a teacher.’

The semantic description of this Roy gives is to say that B’s response is a non-dense predicate predicated of an object, John. A non-dense predicate means, here, that the eventuality of being a teacher in a relevant interval (which might here be the interval corresponding with the typical length of a person’s career) does not need to hold for all subintervals of the interval. In other words, gaps are allowed, times during John’s career when he is not actually teaching (say in the evening when he’s watching TV), and yet it is still true of John that he is a teacher, by way of a characterisation of him.

One thing I’ll say about this immediately is that what it really amounts to is a way of referring with a nominal predicate. It is essentially about the spatio-temporal properties of a nominal predicate. Nominal predicates can be referred to densely or non-densely, and this distinction is essentially what is also at the root of the mass–count distinction in the realm of arguments. For example, mass nouns are ‘divisive’ in the same way that dense predicates are (and as distinct from non-divisive/non-dense predicates); subparts of the mass noun *water* still mean water (Borer, 2005a). As such, I believe that Chomsky would rule this (along

with all acts of referring) as part of the *use* of I-language. All acts of referring are uses of I-language for Chomsky, as we saw above, section 6.3.1, they are ‘how it happens to get used sometimes’. In this light, Roy’s further correlation with syntactic description is interesting. For Roy, there is a ‘very tight connection between semantics and syntax’ (Roy, 2013, 91); she doesn’t, I don’t think, go as far as other syntacticians in claiming that meaning distinctions *follow from* syntactic ones, but her work is certainly consistent with this programme. But this would create a tension with Chomsky’s view of reference, because it would claim that the way we refer is a property of I-language, whilst Chomsky views it as a *use* of I-language.¹³

For now, then, I’ll assume that with respect to non-dense predication, Roy is talking about a function of language, *the way we refer* to nominal predicates. Now, Roy herself shows that there is a certain cross-linguistic consistency in viewing ‘characterisations’ as non-dense predicates. However, she doesn’t hold that there is a one-to-one relation between characterisation and non-dense predication because she also shows that modern Irish (or ‘Irish Gaelic’) does not have characterising predications made out of non-dense predicates (see 1.2.2 for the details of characterisations as non-dense predications). But Roy doesn’t conclude that Irish simply lacks characterising predications; she rather argues that it achieves the characterising function with a different syntax–semantics. If one and the same function can be performed by two different flanker functions (or two different syntactic–semantic descriptions in Roy’s terms), then this is clearly—on its own—a knock-down argument for view (1) above, the view that a sentence function can be deflated into the flanker properties. The evidence for this is presented below. However, the right conclusion I think is not to consider that one sentence function can be fulfilled by two different flanker functions (or flanker syntax–semantics in Roy’s terms), but rather that there are essentially two different meanings here, albeit which are inter-translatable between Irish and languages (like French) that make use of non-dense predicates for characterisations. I don’t really justify this position but I elaborate it in chapters 7 and 8.

6.5.1.1 Characterising in Irish

How, then, is characterising done in Irish? Irish has two copulas, *is* and *bí* which—like Spanish—have generally tended to be correlated with permanent/essential/inherent/i-level qualities (*is*), or temporal/s-level qualities (*bí*). However, the comparison with Spanish is not exact. For example, where Spanish *ser* (the copula apparently correlating with Irish *is*) goes with non-dense predicates (classifier phrases in syntactic terms), this is ruled out for

¹³There may be some wiggle room here if Chomsky were to say that ‘the fact of referring’ at all was a use of language where the ‘way we refer’ is governed by syntax. But I don’t see clearly how this would make sense.

Irish *is*. In (9), there is a lifetime effect (Sean must be dead), a sign that the predicate is maximal, not non-dense:

- (9) Ba dhochtúir Sean. Irish
 IS.PAST doctor Sean
 ‘Sean was a doctor’.

Characterising sentences in Irish are constructed with the other copula: *bí*.

- (10) Tá sé tinn. Irish
 BÍ.PRES him sick
 ‘He is sick’.

In fact (10) is ambiguous between a reading where the predicate ‘sick’ is understood as applying in a temporary way, e.g. after someone had just caught a cold, or in a more general way, as in the way that one might describe someone as having a ‘sick’ character (i.e. ‘I can’t believe you found that person’s misfortune amusing, you’re sick’). This in turn can be described, respectively, as a difference between a dense interpretation of the predicate (where it holds of the subject continuously), and a non-dense one, where there can be ‘gaps’ in the predicate’s holding of a subject within a given interval. And this in turn can allow us to say that we are dealing, respectively, with either a situation-description or a characterisation.

However, there is a strong reason to believe that the characterising reading of (10) is not in fact due to a non-dense predication. The *bí* copula is generally understood not to be compatible with nominal predicates. Recall that we established in chapter 1 that non-dense predicates were classifier phrases, therefore essentially nominal in character. The incompatibility of *bí* with nominals can be seen by the ungrammaticality of (11) (contrast with (9)):

- (11) *Tá sé dochtúir. Irish
 BÍ.PRES he doctor
 intended: ‘He is a doctor.’

The only way a nominal predicate can be used with *bí* is if it is introduced by a preposition:

- (12) Tá sé *(ina) dochtúir. Irish
 BÍ.PRES he in.AGR doctor
 ‘He is a doctor.’

This creates a predication which is not maximal, as the absence of lifetime effects attests when it is in the past (contrast with (9)):

- (13) Bhí Seán ina dochtuúr. Irish
 BÍ.PAST Sean in.AGR doctor
 ‘Sean was a doctor.’

So we are stuck: We can’t simply claim that sentences like (13) are non-dense predications, because we have seen that non-density is a property of the predicate being a classifier phrase, a nominal. However, here we have a prepositional phrase. In (10) we had an adjective. Furthermore, we can’t argue that these apparent prepositional and adjective phrases are in reality nominals because we have seen evidence that the copula *bí* doesn’t go with nominals. How then can we analyse examples like (10) and (13), which appear to be the Irish way of doing characterisation?

The starting observation is that Irish sentences with *bí* allow for generic/habitual readings. This is interestingly not allowed for Spanish sentences with that apparent *bí* correlate *estar*. Compare:

- (14) Tha Calum a’smocadh ’s a’chidsin. Irish
 BÍ.PRES Calum at.smoke.VN in the.kitchen
 ‘Calum smokes in the kitchen.’/‘Calum is smoking in the kitchen.’
- (15) Juan está fumando en la cocina. Spanish
 Juan ESTAR.3SG smoking in the kitchen
 ‘#Juan smokes in the kitchen.’/‘Juan is smoking in the kitchen.’

(The ‘#’ sign of (15) indicates that ‘Juan smokes in the kitchen’ is an impossible reading.) The ambiguity of (10) and (14), then, is claimed to be an effect of the possibility of existentially or generically quantifying the eventuality variable of the sentence. The characterising reading then is in fact a generically quantified dense predication, and not a non-dense predication, as was argued to be the case for French, Spanish, and Russian. (10)—in its characterising reading—could then be formalised along the lines of:

- (16) GEN e : sick(e) & Holder(e , he)
 it is habitually true of the relevant eventualities e that they are eventualities of ‘he’ being sick

A piece of evidence that Roy puts forward to support this analysis is this: given what we’ve seen so far, a possible generalisation capturing which languages will have ‘true’ non-dense adjectives will be predictable from whether the language generally allows adjectives to act as nouns without any further nominal marking of the adjective (called ‘nominal ellipsis’). Certain languages allow this, so, for example, in Spanish one can say:

- (17) Quiero comprar la roja/la barrata/la inglesa. Spanish
 want.1SG buy the red/the cheap/the English
 ‘I want to buy the red/cheap/English one.’

As we can see from (17), Spanish allows an adjective like *roja* ‘red’ to function as a noun without requiring any kind of nominal marking (as in the ‘one’ in the translation that seems to be necessary in English). So languages that, like Spanish, allow bare adjectives to function as nouns should also be the languages that allow adjectives to be non-dense characterisations. This is clearly the case for Spanish, as we have seen. The flip-side of this generalisation is that if a language doesn’t let its adjectives function as non-dense predicates, then we should expect that it also doesn’t allow them generally to act as nominals without further nominal marking. Is this true of Irish? Yes, as (18) shows:

- (18) Tá meag iarraidh an *(cheann) dearg. Irish
 BÍ.PRES me at.seeking the thing red
 ‘I am looking for the red thing/one.’

6.5.1.2 What does this mean?

I think this clearly shows that any attempt to deflate a sentence function into properties of its flankers is doomed. This includes the view that the flanker properties are themselves functions (as Higgins believed, and which is essentially the view I take in chapter 7), but it also crucially includes the view that syntax determines sentence meanings, if we think of these flanker properties (as Roy does) as ultimately matters of syntactic categories. As I said before giving Roy’s Irish analysis, her own conclusion simply seems to be that generic dense predications are another way to do characterisation (so she isn’t herself a syntactic determinist it seems). But then she also seems to view—along with Higgins—sentence functions in terms of their flankers (so there does at least appear to be an internal tension in her analysis).

Now, there could be an interesting way of justifying a looser relation between sentence functions and flanker properties. There seems to be a strong inclination to say that the following sentences have the same function of being characterisations:

- (19) Bhí Seán ina dochtuúr. Irish
 BÍ.PAST Sean in.AGR doctor
 ‘Sean was a doctor.’
- (20) Jean était médecin. French
 John was doctor
 ‘Jean was a doctor.’

Where does this strong inclination come from? My assumption is that it is because these two sentences would appear to be perfectly acceptable translations of each other. ‘Inter-translatable’, then, could be a way (an admittedly question-begging way) of avoiding talking about a particular shared function for sentences of different languages (or indeed different sentences within one language), whilst still expressing an essential similarity between them. The way I think this could be rendered more interesting (and less question-begging in so far as it would be trying to answer the questions begged by the notion of inter-translatability) would be to conduct a research programme whereby sentences like (20) are translated into as many different languages as possible, precisely in order to see what variation is possible and what variation is not possible. This could lead to interesting conclusions about a potential sentence function and what the flanker properties can and cannot be.

My own conclusion isn’t quite this. I think that we should conclude that (19) and (20) really do have different meanings. They are not the ‘same’ sentence simply translated into a different language. However, they are also inter-translatable. This means that sentences with different meanings can be used in the same contexts in different languages. I don’t see why this conclusion should be impossible. And it frees us to simply deal in flanker functions and allow for as many different types of sentences as there are different types of flanker function.

6.6 Conclusion

This chapter is the first of three chapters which develop the approach to classifying copular sentences which I advocate in this thesis. Having established that the logical notion of predication and logical–metaphysical notion of identity cannot be used to exhaustively *explain* the taxonomy, this chapter has gone back to square one, so to speak, to try and understand what Higgins considered his taxonomy to be a taxonomy of. We looked in more detail at the functional definitions of Higgins’s copular sentences and copular sentence flankers, 6.2. We then went more deeply into Chomsky’s notion of I-language, the biological endowment which gives rise to the human ‘linguistic code’ (in Scott-Phillips’s terms) and Scott-Phillips’s notion of ostensive communication, which is also ultimately a biological endowment (a recursive theory of mind) allowing us to mentally manipulate (and read) other minds, a form of extended social navigation, 6.3 and 6.4. Words, sentences, and referring are all uses of I-language for Chomsky, and Scott-Phillips would agree: a sentence is an instance of mentally manipulating one’s audience using the linguistic code (remember we don’t have to use linguistic codes/I-language in order to ostensively communicate). This notion is very general, and doesn’t take us far in distinguishing types of copular sentence. Furthermore, in

common with other human actions, it is an exercise of free will and so not susceptible to a true scientific explanation. The best we can hope for is a better *description* of human behaviour in this domain, 6.5. Conceptual difficulties remain, the most notable of which is the relationship between sentence functions and flanker functions. One thing we can say, though, is that we can't deflate sentence functions into flanker functions: 6.5.1 showed that Roy's 'characterisation' sentence function can be fulfilled by both non-dense predicates (in e.g. French) and generic dense predicates (in e.g. Irish). My own view is that we should drop talk of sentence functions. French's non-dense predicates and Irish's generic dense predicates really do have different meanings; it just so happens that they can be used in the same contexts, that they are inter-translatable. I don't justify this position directly, but in the next two chapters I elaborate my approach to the copular taxonomy in a way which is consistent with it.

Chapter 7

Pieces of the puzzle

7.1 Introduction

This chapter looks at some of the conclusions that we can draw from the discussion so far in this thesis which will be of use in an eventual description of copular sentence meanings. It is therefore necessarily a slightly eclectic chapter. I'll look back over the previous chapters and highlight what I think from them usefully enters into a full description of the distinct meanings of copular sentence types. I'll deal with some outstanding issues which I've promised to do, and introduce some new factors that haven't been properly discussed so far, but which I think are essential to an eventual description. The coherence of the chapter resides in the way that it leads up to the final section 7.7.2, where I try to exhaustively list all the factors of description that I have discussed and indicate which ones I think can be used to distinguish types of copular sentence.

I start, in 7.2 by summarizing (very briefly) what I think we can take from the discussion of chapters 2 and 3, before turning to one of the loose ends which I've mentioned already at various points—the potential role of a metaphysical ontology, 7.3. 7.4 briefly mentions the potential utility of 'On Sense and Reference' for understanding identity statements, before I discuss, in 7.5, what I think we can take from syntax. Without syntax there would certainly be no copular sentences. However, the discussion in the last chapter has given us perhaps a rather more austere notion of what is truly syntactic, truly structural. On this basis, syntax doesn't really do much work in distinguishing copular sentence types, understood as a categorisation of their meanings. It would be rather like relying on the fact of plant cell division to explain the variety of forms of plant life. Without plant cell division, there would be no plants in the first place, but once it is in place, we need to resort to other factors (amongst which—though not exclusively as Darwin already recognised¹—natural

¹I am convinced that Natural Selection has been the main *but not exclusive* means of modification.'

selection). 7.6 relates to a notion that we have encountered in passing but not officially named thus far—information structure—and introduces some basic ideas which will be of use in the next chapter. 7.7 collates the factors of description into a list. It also makes the important point that there is not one taxonomy of copular sentences (any more than there is only one taxonomy of any natural world phenomenon generally); taxonomies reflect theoreticians' interests. If, for example, one is interested in the property of inversion of the flankers, then one can exhaustively classify copular sentences in this respect (and this would, I think, be one way we could talk about a truly syntactic taxonomy, indeed explanation, of copular sentences). I think, though, that we can talk about a 'full description' which is the aggregate of all the factors that enter into making semantic distinctions between copular sentences. These factors relate to the flankers, not the copula or any kind of relation between the factors, and the various permutations of these 'flanker properties' give rise to the full variety of copular sentences.

7.2 What logic gives us

Chapter 2 didn't really give us anything from predicate logic that we can use to distinguish one copular sentence from another. The notion of a predicate as a Fregean function is (1) not part of the semantics of logic, so much as its combinatorics (and therefore not relevant if we assume that the copular taxonomy is of meanings) (2) uniform to the entire taxonomy anyway, so cannot serve to distinguish copular sentences.

Syntax makes use of Fregean-style functions, as we saw in chapter 2, for theta roles, argument structure, and type-shifting (more of a semantic than syntactic function) *inter alia*. Even the core combinatorial operation of syntax—Merge—can be viewed as a function. But I would hesitate to say that this means that logic gives us the notions of a theta role, argument structure, or the way a given lexical item can refer to a property or an object. Functions are simply parts of a scientific theory. As McGilvray (Chomsky & McGilvray, 2012, 167) puts it, scientific theories generally are 'statements of functions aimed at describing and explaining what there 'is,' where it is assumed that there 'is' something 'there' that can be captured by a theory, and is captured by the correct theory'; these statements are what he calls 'formal' functions. However we must bear in mind that, 'formal functions themselves thought of as sets of symbols and their theory-specified forms and specified allowed combinations are invented 'objects,' not natural ones. They amount to the 'syntax' of a formal symbol system, and those who are adept in the relevant formal 'language' apply the system's symbols in a regimented way' (168). Functions are 'formal tools', 'artifacts'.

(Darwin, 1859, 6)

Following on from this, I wouldn't even wish to say that the phenomenon of predication, in so far as it is a product of the natural world in some way is literally Fregean functional application. Functional application is just a description of the fact that e.g. putting *John* and *is happy* together gives you something which is true or false. But it doesn't say naturalistically how or why this happens. But Merge is, for Chomsky, an object of the natural world. In 7.5 we'll come back to Merge and say more about it and about the way that it is the essential naturalistic notion behind functional application, a necessary component, at least, of predication. Logic, then, is really of no use to us as any kind of explanatory primitive.

7.3 A note about metaphysics and property ascription

At several points earlier in the thesis, I have promised to say what role metaphysics might play in explicating copular sentence meanings, notably with respect to the notion of 'property ascription'. Chapter 2 scouted logic and syntax for a meaningful notion of predication and concluded that logical predication (i.e. functional application) was not meaningful, and that it wasn't obvious that it was possible to reduce 'predicate' to syntactic terms; there were empirical problems with 'the constituent that assigns a theta role to an external argument'. Throughout that chapter we referred to 'property ascription', the notion that is standardly taken by linguists as the meaning of predication.

An obvious place to look for a meaningful notion of property ascription is metaphysics. This, after all, is the field which contains properties as primitive parts of mind-external reality. The flankers of copular sentences—or at least some of them—presumably have extensions. Extensions are the denotations of linguistic elements, the things in the world that linguistic utterances refer to. And generally, if we have a categorisation scheme of possible mind-external entities and how they relate to each other, then what we say, in so far as we refer to the world, will surely be delimited and explained by this scheme. Perhaps Aristotle's *Categories* (Aristotle, 1963) is the most well-known attempt to exhaustively elucidate mind-external reality, and it continues to be influential today as the basis of contemporary categorisation schemes, e.g. (Lowe, 2009; Lowe, 2013).

So metaphysics may provide a basis for the meaning of (at least part of) some copular sentences. Let's focus on the notion of 'property ascription'. What, metaphysically, speaking might a property be?

It seems there are two ways to go here. The first way is to simply correlate 'property' with one or more of the categories of a metaphysical ontology. Aristotle has what can be given as a four-category ontology:

(1)

Non-individual substance	Non-individual non-substance
individual substance	Individual non-substance

There are two distinctions: (1) Individual vs. non-individual. This is sometimes called the ‘particular’ vs. ‘universal’ distinction. If the category is the substance *horse*, then an individual horse is e.g. my pet horse Dobbin. The non-individual horse is just the category ‘horse’, which might be expressed in English with *a horse* (in *Dobbin is a horse*) or *horses* (in *Dobbin and Zedanzig are horses*). (2) Substance vs. non-substance. Overall, Aristotle lists ten categories in the *Categories*. One is the substance category, the others are ‘non-substance’ categories. The following table from Smith (2017) gives the categories with examples:²

(2)

Category	Examples
Substance	<i>Socrates, man, horse</i>
Quantity	<i>four-foot, five-foot</i>
Quality	<i>white, grammatical</i>
Relation	<i>double, half, larger</i>
Location	<i>in the Lyceum, in the marketplace</i>
Time	<i>yesterday, last year</i>
Position	<i>lies, sits</i>
Habit (lit. ‘having, possession’)	<i>has-shoes-on, is armed</i>
Action	<i>cuts, burns</i>
Passion (lit. ‘undergoing’)	<i>is cut, is burned</i>

From this, what might we correlate with a property? Traditionally ‘property’ has been the label given to universals, which in the Aristotelian scheme correspond to ‘non-individual substances’ and ‘non-individual non-substances’. In that case, in terms of the list of categories above, we get properties coming out as any of the examples in (2), with the exception of *Socrates*, who is an individual substance. The other examples of substance are non-individuals, and all the non-substance category examples are non-individuals. An example of an individual non-substance would be e.g. my handkerchief’s whiteness, Reese Witherspoon’s five-footedness.³

Of course ‘property *ascription*’ is suggestive of the *relation* between a property and something else. There are in fact exactly two relations between metaphysical categories

²Is the list exhaustive? Are these ten categories all the entities that exist and the only ones? About this Cohen (2016) says, ‘Although Aristotle never says so, it is tempting to suppose that these categories are mutually exclusive and jointly exhaustive of the things there are’.

³For many categories it isn’t obvious to me what an example would be e.g. What is the particular/individual version of *yesterday* (*today’s yesterday*)? Metaphysical works tend to focus on examples of ‘qualities’.

for Aristotle in the *Categories*: ‘saying of’, and ‘being in’.⁴ ‘Saying of’ is essentially the superordinate–subordinate relationship between levels of a hierarchy within a substance or non-substance. So we can say of Dobbin that he is a horse, we can say of horses that they are mammals etc.⁵ ‘Being in’ is to do with the way that a category cannot exist separately from what it is in.⁶ Strictly speaking, for Aristotle, this only related individuals of the substance and non-substance category. So Callias’s generosity (an individual quality) can be in Callias (an individual substance). But it seems standard to say that e.g. Callias’s generosity being in Callias is just the literal version of what we would actually say i.e. *Callias is generous*.⁷

An example of a neo-Aristotelian four-category ontology is given in figure 7.1. For Lowe, then, the property categories are his ‘kinds’ and ‘attributes’. A sentence like *Dobbin is a horse*, then, is an ‘instantiation’ relation between a kind (horse) and an object (Dobbin). *Dobbin is white* would be an ‘exemplification’ of whiteness (an attribute) by Dobbin. Exemplification isn’t a primitive relation for Lowe, but rather a consequence of going round the square either via kinds or modes. The latter way is consistent with what I said in the last paragraph about the way *Dobbin is white* should literally be thought of as Dobbin’s whiteness is in Dobbin (or Dobbin’s whiteness (a mode) is characterised by Dobbin (an object) in Lowe’s scheme). An instantiation of an attribute by a mode might be e.g. Dobbin’s whiteness instantiating whiteness.⁸

⁴There is some debate about how exactly these are to be interpreted, though Ackrill is explicit that they are metaphysical relations:

It is often held that ‘said of’ and ‘in’ introduce notions of radically different types, the former being linguistic or grammatical, the latter metaphysical or ontological; and that, correspondingly, the word translated ‘subject’ (literally, ‘what underlies’) means ‘grammatical subject’ in the phrase ‘said of a subject’ and ‘substrate’ in ‘in a subject’. In fact, however, it is perfectly clear that Aristotle’s fourfold classification is a classification of things and not of names. . . (Aristotle, 1963, 75)

⁵What, in fact, the substance would then be (‘horse’, ‘mammal’, ‘animal’?) I am unclear about.

⁶Lowe (2013, 56) characterises this as *constitution* (i.e. if A is in B, then B is constituted—at least partly—of A).

⁷I take Lowe (2013, 57) to be making this point (Lowe uses the term ‘inherence’ for ‘being in’):

When, for instance, we say that Dobbin *is white*, we are making no reference to his *individual* whiteness, even if it is *because* this individual whiteness ‘inheres’ in him that whiteness (the universal) is predicable of him.

⁸Whilst apparently entirely unrelated, it’s interesting to note that the ‘functional’ classification that Dik (1983, 131) presents is closer to a metaphysical scheme than most other classifications. Dik has five ‘semantic relations’ involving the copula:

- (3) a. Property assignment = *John is clever* b. Class membership = *John is a man* c. Class inclusion = *An elephant is an animal* d. Identity = *John is the man I need* e. Location = *John is in the garden*

The differentiation between ‘properties’ and ‘classes’ in particular is reminiscent of Lowe (2009)’s neo-

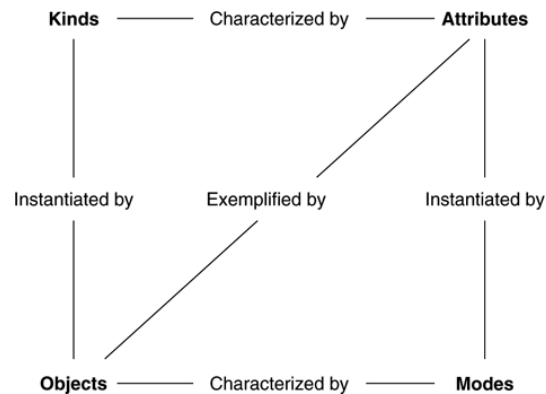


Figure 7.1: A four-category ontology, from Lowe (2006a)

I'll say one more thing about the metaphysics of properties. A distinction is often drawn between property 'in the strict sense' and 'general accidents'. The former is a property that is necessarily shared by 'all primary substances of the same species' (Lowe, 2013, 56), i.e. all individuals of a substance, in Aristotelian terms. This is the case of warm-bloodedness (a non-individual quality) being in Dobbin. It is a property in the strict sense because all individual horses are warm-blooded. (I assume that this 'strict property' status would also apply between warm-bloodedness and the kind 'horse', though Lowe doesn't explicitly give an example of this). Otherwise, where a property does not hold of all the individuals of a substance, we talk about 'general accidents'. So another non-individual quality, whiteness, can be in Dobbin. But not all horses are white, so this is not a property of Dobbin in the strict sense.

So we have two definitions property ascription from metaphysics. A broad one that aligns them with all the universals (substance and non-substance) in Aristotle's schemes, and the (saying of or being in) relations that hold between them. And a narrower one of 'strict properties'.

7.3.1 Returning to language

I've made this slight philosophical detour partly because I think it is important to highlight that 'property' is not quite an 'all-purpose' term that we can simply apply to any predicate. Not in metaphysics in any case, and if not in metaphysics, then I don't know where else to look (as logic was not useful to understand a substantive notion of 'property', as we saw in chapter 2). Now clearly metaphysics isn't *that* restrictive, if we take the broad notion

Aristotelian 'attribute' and 'kind' classes respectively, what Aristotle would have called 'non-individual substances' and 'non-individual non-substances'. Dik's class inclusion is a 'saying of' relation between non-individual substances, in Aristotelian terms.

of property. Aristotle's list of categories is quite comprehensive and many of them, as the reader can verify, lend themselves to being expressed as the linguistic predicate of a copular sentence.

But not everything can be a property according to the metaphysical schemes I've outlined. Individual substances (objects) for a start (though this is also taken care of in logic). For Aristotle everything else depends on individual substances (which are thereby 'fundamental' ontologically), but individual substances do not depend on anything else. Specifically, they are not 'said of' or 'in' anything else. What about specificational sentences? What about *the culprit* in *The culprit is Brian*? Well, hopefully by this point in this metaphysical detour, it should be clear that we are going to struggle to place *the culprit* among the categories. The temptation is to say that *the culprit* is an object. But nobody makes this claim exactly, for the simple reason that we can't actually pick out which object in the world *the culprit* is. If we could, then of course Brian and the referent of *the culprit* would be numerically identical metaphysically, and there would be a relation between them that we could express with an identity statement (which would have the meaning of an identity statement, such as we encountered in chapter 3). But there is no place in Aristotle's categories for objects with the 'obscure kind of referentiality' (as Higgins (1979, 216) put it) of *the culprit*. In fact, as I'll argue just below, it is essentially just Donnellan's (1966) attributive kind of reference, which I think can be innocently thought of as referring to 'putative' objects, objects that must exist but which we aren't in fact acquainted with. But Aristotle's categories, and the wider field of contemporary metaphysical ontology doesn't include putative objects; it only includes actual objects. As such, *the culprit* in the *The culprit is Brian* cannot be said to be a property ascribed to an object, if we are thinking in metaphysical terms.

7.3.2 Identificationals

Identificationals are perhaps the class of copular sentence that has attracted the interest of metaphysical description more than any other. This is perhaps due to the philosophical interest in demonstratives in the work of Kaplan (1989); Dummett (1981) describes *this* as 'a mere chunk of an amorphous reality'. Moltmann (2013) has recently argued that the demonstrative in an identificational refers in some way to a 'trope' (an individual non-substance in Aristotle's terminology, a 'mode' in Lowe's, see figure 7.1) which she takes to be the objects of direct perceptions (as Lowe (2008) argued). Moltmann (2013) argues that the demonstrative in identificationals (which she calls a 'presentational pronoun') is not a normally referring expression and introduces a novel compositional semantics for it involving the use of a function 'mapping an epistemically possible or conceivable world to the entity

that according to that world is the bearer of the trope in question' (57). I don't think the arguments she raises to justify this position are convincing, but this is not the place to review those details (though see Woodard (2013a)). Heller and Wolter (2008) bring in the notions of 'quiddities' or 'essences' into their analyses of identificational. Saying that *This is a camera* can (in the right context) tell us something about the essence of an object; this is ultimately understood, it seems, as a case of sortal predication (a non-individual substance being said of an object in Aristotle's terms, or a kind instantiation in Lowe's, figure 7.1).

7.4 What (else) philosophy gives us

So metaphysics gives us a notion of property ascription which is quite broad, but not so broad that it would extend to classifying specificational sentences as property ascriptions. We also encountered metaphysics at the start of chapter 3, as we looked for a meaning for identity statements that might restrict how we could categorise copular sentences. We concluded that the 'pure' (logical or metaphysical) notion of identity amounts to a general statement of the numerical identity of any object. Particular statements of this fact are always tautologies. Informative identity statements of natural language are not tautologies, therefore they are not expressions of numerical identity. So metaphysics doesn't really help us with identity.

We then looked at some more ideas of Frege about the meanings of identity statements. Generally, works like 'On Sense and Reference' are included in the 'philosophy of language' canon. Is it possible to reduce Frege's meanings of identity statements to logic or metaphysics? The 'On Sense and Reference' theory could be viewed as a statement of metaphysical facts. Remember that an object's 'sense' was for Frege a mind-external aspect of it. And, of course, the object would be too. There is an important relation between names and senses as well, and in so far as names are *not* mind-external objects, I think we have to conclude that the 'On Sense and Reference' view of identity statements is a mixed theory of metaphysical categories (objects and aspects of objects) and a mental one (names).

A main conclusion from the discussion of identity is that it is almost immediately obvious that none of the theories of identity statements from philosophy that we've encountered lend themselves to an account of specificational sentences. Whatever specificational sentences are, claiming that they are kinds of identity statement has no foundation in philosophical discussions of identity statements.

This raises the question of how identity sentences should be understood. I don't offer a worked out analysis of identity statements here, but I do think that sticking with Frege's 'On Sense and Reference' view could be helpful. It may be important to include the notion of the

distinct senses of the flankers in any eventual description; whether this would correspond to anything in a more modern linguistic idiom is unclear to me, however.

7.5 What syntax gives us—syntactic predication

Frege's definition of predication in terms of functional application is often held to be a logical explanation of the meanings of some copular sentences. Chapter 2 and section 7.2 above try to rebuff that claim. But such functions may well provide a formal description of a natural world object. There's no controversy (in linguistics at least) that copular sentences are uniformly 'linguistic' or 'grammatical' predications.⁹ Here, I will say that at least a necessary component of this is a truly syntactic notion, part of Universal Grammar.

Strictly speaking, I don't know if syntax on its own provides us with a notion of predication. Syntax provides the core combinatorial mechanism that puts say *the house* and *blue and indigo* together to form the unit *the house blue and indigo* in either *John painted the house blue and indigo* or *The house is blue and indigo*. But putting things together doesn't necessarily go beyond just that, the creation of a set. The fact that there is (1) a truth value, (2) a particular kind of meaning, seem to be things that go beyond the mere *putting together* of *the house* and *blue*. By 'a particular kind of meaning' I mean something distinct to what we see in:

- (4) The colours that lie between turquoise and violet are blue and indigo.

In (4) *is blue and indigo* doesn't obviously mean the same thing as it does in *The house is blue and indigo*. Only the latter is ascribing the properties of blue and indigo to the house; (4) is listing colours. But in both cases there is a truth value.¹⁰ And whilst I think that we can differentiate between property ascription and listing using some of the properties I list below in section 7.7 (notably the information structural ones), why putting *the house* and *blue and indigo* together leads to the meaning of property ascription, say, I have no idea. I think the explanation goes beyond syntax alone, though recent accounts by Collins (2011) and Reichard (2013) do precisely endeavour to show that the unity of such sentences—the way they give rise to truth values—is a product of their grammatical structure. The claim I

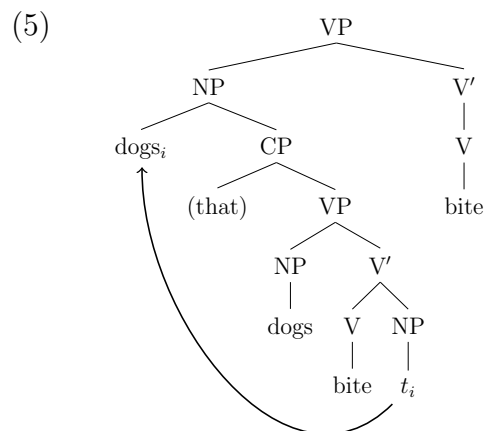
⁹And recall again that all the copular sentences, including equatives, would have been analysed in terms of functional application for Frege, section 2.2.7.

¹⁰As we already noted in the last chapter, 6.2, Higgins actually gets close to denying this for specificational: '[I]t is not obvious that the notion of having a truth value. . . is pertinent to lists, for one tends to classify lists as correct or incorrect, complete or incomplete, useful or useless, but hardly as true or false' (Higgins, 1979, 214). This certainly isn't a mainstream view, but it's possible that one could interestingly pursue this idea in terms of a notion of 'rectitude' underpinning both true predicational sentences and 'correct' specificationals, in the spirit of Anselm of Canterbury's approach to truth (Anselm of Canterbury, 2000).

would make here is weaker; I think syntactic combination is necessary for sentential unity—and indeed the meaning of predication if there is one—but I doubt that it is sufficient. In what follows, I'll just give a sketch of the nature of syntactic combination that I think is a necessary component of predication.

7.5.1 Merge

What, then, is the syntactic mechanism that helps give rise to predication? Well, for this I'll rely on the combinatorial mechanisms that Universal Grammar makes available. The central operation is 'Merge' e.g. Chomsky (2015) *inter alia*; this is a recursive operation that puts the atomic elements of natural language together such that, for example, the verb *sleep* joins with the present tense morpheme *-s* to create the word *sleeps*, and *sleeps* joins with the noun *John* to create the sentence *John sleeps*.¹¹ This of course simplifies and abstracts from any underlying/invisible structure, but the basic merging principle is the same. As we just saw, the operation is recursive: it applies to its own output; having applied to *sleep* and *-s* to form the output *sleeps*, it is then applied to this output to form *John sleeps*. And in theory this process can go on—like any recursive process—indefinitely, although in practice there are 'comprehension' limits: Pinker (1994) gives the example *Dogs dogs bite bite* which is parsable without too much effort, meaning 'Dogs that other dogs bite themselves bite' or 'Dogs that are bitten by other dogs are themselves biters' (compare *Dogs that owners stroke don't bite*). This could be given in a tree (this assumes a version of Kayne's (1994) raising analysis of relative clauses):

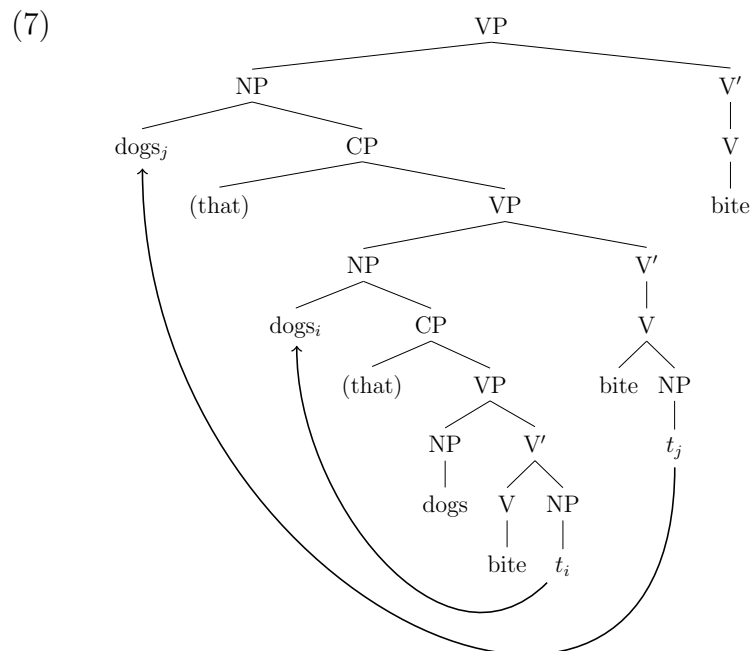


¹¹Not all linguists accept that word-level (morphological) structure makes use of the same mechanisms as sentence-level (syntactic) structure, but it is quite a standard view in UG; 'Distributed Morphology', in particular, is an influential framework that holds that the same structural operations that apply to syntax apply to morphology (Halle & Marantz, 1993; Halle & Marantz, 1994).

Merely extending this by one further level of embedding, however, completely bamboozles the human parser:

(6) *Dogs dogs dogs bite bite bite.

This would have the tree:



Note that a structurally comparable sentence such as

(8) ?Dogs strangers their owners like stroke don't bite.

whilst still not straightforward to parse, is certainly an improvement on (6) and seems just about comprehensible with the right intonation. So the mere fact of embedding the same material within itself also seems to account for the difficulty in parsing these sentences.

So recursive structure building can go on indefinitely in theory but meets comprehension difficulties in practice (possibly to do with human memory limitations). The trees above contain several 'branchings'. Each such branching represents an instance of Merge: *dogs* to *bite*, the subsequent Merge of *bite dogs* to *dogs* to form *dogs bite dogs* in (7) (before raising of the object to leave the trace t_j as part of the relative clause formation). The arrows represent instances of 'move' (or 'Internal Merge', Chomsky (2015)), which as we said in

section 2.3.7 is effectively the same operation as Merge, just with the merged item coming from an already-built portion of the tree, as opposed to being merged from somewhere external to the tree.

This is the basic building block of syntax, and is at the heart of how the flankers of a copular sentence are joined together. There are many technicalities to explicating Merge further: whilst Merge is certainly simple, it isn't quite as simple as mere set formation. Notably, the trees above demonstrate the importance of 'headedness' or the 'projection' of the label of one of the merged elements. We see in (7) for example that the merging of a verb *bite* to a noun *dogs* produces a verbal structure (labelled V'), not a nominal structure. Why this should be the case is a long-standing technical problem in Universal Grammar (Chomsky, 2013; Chomsky, 2015).¹²

7.5.2 Phases

I introduced the notion of a 'phase' in chapter 5, section 5.2.2.3.2. They are essentially measurements of the structural distance that applies to various syntactic operations. We may say that the 'tense cycle' is closed off by the phase head, C. Hinzen (2012b) identifies exactly three such cycles in grammar, the nominal cycle (with upper limit/phasal head D), the verbal cycle (with upper limit/phasal head transitive verb phrase, v^*P , see Chomsky (2008)), and the tense cycle (with phase head C), ordered in such a way that the nominal cycle embeds in the verbal cycle which embeds in the tense cycle. On the basis of the arguments in section 4.4, I take the verbal cycle to be inapplicable to copular sentences, so we just have a nominal cycle embedded in a tense cycle.

The crucial, albeit frankly trivial, point about cycles is that they are intervals of coincidence (temporal coincidence seems most likely). More than one thing can occur in them. To say that two elements coincide is ultimately to state: two elements are situated spatially or temporally within one spatial or temporal interval. This doesn't make the elements themselves identical. Copular sentences, then, comprise unique intervals in which both flankers reside. The different kinds of copular sentences, along the lines of Higgins's classification,

¹²The solution to the problem of which head projects and why may also provide an account of predication, in the sense of the difference between attributive and predicative modification. In *red jumper* for example *red* modifies *jumper*. But why should this be so? (1) Why should either element modify the other? (2) Given that one does, why should it be *red* and not *jumper*? Syntactically, it is *jumper* that projects its head in *red jumper*, which is thus a noun phrase, and there is attributive modification of *jumper* by *red*; the constituent that projects is the one that gets modified. This is at the phrase level. At the sentence level, things seem to be the other way round. In copular sentences like *The jumper is red*, we are clear that *red* still modifies/is predicated of *jumper*, not the other way round, though the sentence is not a noun phrase; rather there is predicative modification of *the jumper* by *red* here. As we saw in 2.3.4, for many syntacticians, there is a 'core' predication phrase in *The jumper is red* where *red*—as the predicate—is essentially what projects (Bowers, 1993; Roy, 2013), exactly the opposite of what we see with *red jumper*.

are to do with the different kinds of flanker which reside in the intervals, which are cycles. The cycles themselves are the same for all copular sentences.

I take there to be empirically three main phrasal categories that encompass *both* the flankers of a copular sentence (i.e. which are part of Hinzen’s ‘tense cycle’ referred to just above): aspect, tense, and the clausal category C, which itself may be expanded into further phrasal categories, to do with modal force. Cross-linguistically, mood, tense, and aspect tend to be morphologically in the order mood before tense before aspect pre-verbally, and the mirror image post-verbally (i.e. aspect > tense > mood), suggesting an ‘order’ in which these phrasal categories are built whereby aspect precedes tense, which precedes mood, the final category to be added (though see Cinque (2014) for a substantial refinement of this generalisation). Whether it really makes sense to consider these three phrases as forming part of one cycle (i.e. the tense cycle) is not clear to me at present. It would make sense to assume that the first one to encompass both flankers should serve as the relevant interval for identifying the exact point that they initially come together, i.e. aspect, though it is somewhat doubtful to me that all copular sentences do in fact have aspect, and this has certainly been doubted for nominal sentences (as we saw in section 4.2.2).

But let us just assume that all copular sentences are tense/aspect coincidences. The notion of syntactic predication can then just be stated as this fact of phrasal, or cyclic/phasal, coincidence. The flanker-based taxonomy of copular sentences is then a matter of identifying the nature of the flankers, and there should be as many kinds of copular sentence as there are kinds of flanker.

7.5.3 Merge and phasal coincidence—a redundancy

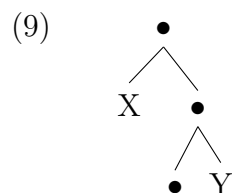
So far, then, we have discussed two formalisms of contemporary Universal Grammar: Merge and phases. But I think that laying them out like this makes it rather obvious that there is an apparent redundancy: syntactic objects are getting ‘combined’ in two different ways: (1) via a sort of ‘chemical bonding’ of the syntactic objects—Merge, (2) via the spatio-temporal coincidence of syntactic objects in a unique interval. It isn’t clear to me that both are needed for my purposes. My preference is to select phases, given they are needed independently to explain the locality of syntactic relations. For the rest of the thesis, I’ll continue to represent syntax using trees, whose branching structure is visually suited to representing merged structures. But all that matters for the purposes of the syntactic composition that syntactic predication is built upon is that syntactic elements are embedded in phases. Can we go any further in terms of understanding the biological basis of predication? I think we can; I think we can think of the phases of syntactic theory as—essentially—the format

of a specific kind of working memory workspace, in the sense of Baddelley (2007). This then leads to the intriguing conclusion that predication is enabled by a particular kind of working memory. This isn't the place to develop this speculation, but see Woodard (2013b) for further details.

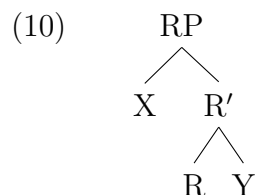
7.5.4 Predication Phrase?

As mentioned at various points in the thesis so far, many authors have sought to align the notion of predication with a syntactic 'Predication Phrase' (Bowers, 1993; Svenonius, 1994; Adger & Ramchand, 2003). It goes without saying that this is rather obviously circular in terms of actually *explaining* what predication is, so I won't really spend much time discussing the rationale for this theoretical construct (which isn't really given a rationale anywhere anyway; it simply seems to serve a descriptive purpose, for aligning predication structures with other syntactic structures in the absence of any other label for them).

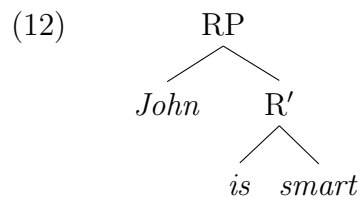
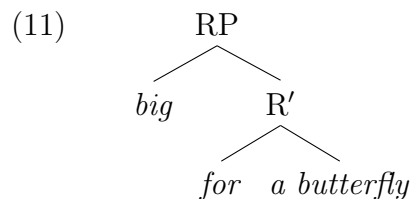
Den Dikken's (2006a) is more in the spirit of how I would view predication here. For den Dikken, predication is not a specific phrasal head, but rather any instance of the structure:



This is simply the configuration whereby Y is predicated of X or X is predicated of Y (the 'non-directionality' is different to the PredP approaches which stipulate that the c-commanded element, Y, would have to be the predicate). For the sake of labelling, den Dikken calls the element that mediates the relationship between a subject and predicate a 'relator', thus the structure of (9) can be viewed as a 'relator phrase':

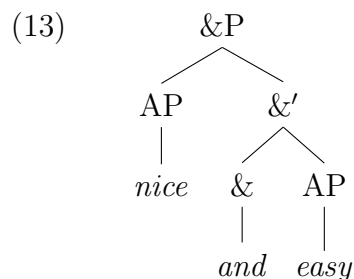


But relator/relator phrase is not a distinct functional category, just an abstraction over many possible instances of predication. This then allows den Dikken to show that many different functional elements can lexicalise the relator. For example, the phrase *big for a butterfly* includes a relator *for* with *big* predicated of *butterfly*, in the same way as *John is smart* has a relator *is* and predication going the other way, *smart* predicated of *John*:



So the relator can be the copula, a prepositional element, ‘...but it can also be T (or Infl), as in *John walks*, or indeed any head that relates a predicate to its subject, including functional heads in the A’ domain of the sentence (such as Topic and Focus)... [W]hat is important to bear in mind... is that “R” is not a designated category; the RP structure in (1) represents a *syntactic configuration* rather than a claim about the lexicon’ (den Dikken, 2006a, 16). Den Dikken says: ‘Any semantic relationship of predication can in principle be translated into a syntactic configuration of the type in [(10)]’. There are two things to say about all this. Firstly, we need to note that the determination of which element is the predicate and which is the subject does not come from the syntactic configuration because of the non-directional nature of the predication relation for den Dikken. Secondly, whilst all predications must be modelled with the configuration of (10), is there an implication the other way: are all configurations of (10) instances of predication in some sense?

My feeling about den Dikken’s view (and the PredP view) is that we shouldn’t be surprised that (10) is the configuration of predication, given that it just seems to be a statement of the X’-schema, which is the foundation of any syntactic structure, according X’ theory (see footnote 9 in chapter 2). Necessarily, in so far as predications are sentences and sentences can be modelled in terms of X’-schemata, then predications will always be modelled in terms of X’-schemata. The second question of the previous paragraph then becomes whether all X’-schemata are predications in some sense. Following Kayne (1994), den Dikken also gives coordination phrases the same structure:



This is good for den Dikken as he considers all predications to involve set intersections (which is complemented with property ascription to give the full semantics of predication). But I expressed some doubts about this in section 2.2.5 of chapter 2. The set intersection of *John* and *kissed Mary* would be John (a thing that is John and a thing that kissed Mary). So whether all X' -schemata are predications, I am unsure about. And whilst (13) clearly expresses set intersection, it isn't obvious that there is any property ascription, which presumably Den Dikken would predict there should be. A safer conclusion seems to me to be that all predications are built on X' -schemata (because they fundamentally require constituents to be merged/phasally coincident). That's a necessary condition of linguistic predication, but not obviously a sufficient one.¹³

So it seems that syntax is absolutely crucial to get predication off the ground. Without the basic capacity to combine lexical items in the first place, there would be no predication. But this is evidently universal to the copular taxonomy. Syntactic predication applies equally to specificationals, predicationals, identities, and identificationals. As such, then, whilst the fundamental role of syntax mustn't be underestimated, it shan't enter into an account of the copular taxonomy, given that it doesn't discriminate the classes.

7.5.5 Is there anything else to say about syntax?

7.5.5.1 Scope relations and the effects of what is composed with what

In terms of what we can say has been truly explained by syntax so far in this thesis, there's only really been one phenomenon, namely the different possible scope readings in sentences like the ones we encountered in section 5.2:

- (14) a. A student seems to be sick (ambiguous: $\exists(a\ student)>seem; seem>\exists(a\ student)$)
 b. A student seems sick $\exists(a\ student)>seem; *seem>\exists(a\ student)$

This is not nothing, but it is precisely the sort of 'meaning' difference that I think is more or less all and only what can be truly explained in purely syntactic terms.

Another example of a truly syntactic account of a meaning distinction would be Cinque's (2010) account of the intersective vs. subsective interpretation of adjectival modification (see Reichard (2013, 121ff.)). Cinque has argued that intersective adjectives are essentially reduced (i.e. unpronounced) relative clauses, their clausal origin accounting for the 'symmetry' between them and the nouns they modify, where subsective modification is a more straight-

¹³See Marelj and Matushansky (2015) and Matushansky (to appear) for an impressive and convincing range of theoretical and empirical arguments against the Predication Phrase/Relator Phrase approach.

forward asymmetry between adjective and noun, where the adjective only modifies the noun, and there is no semantic value that is symmetrically modified by both the adjective and the noun. Reichard (2012; 2016) has also persuasively argued that the sorts of entailments Davidson (2001, 105–22) used as evidence to justify the postulation of event arguments can be captured in terms of configurational relations. Generalising about these cases, we can see that the kind of ‘meaning’ that is captured is to do with what modifies what, what is merged with or phasally coincides with what. As I say, this is not nothing, but it won’t take us very far in distinguishing copular sentences as it is not obvious, where they are concerned, that such scope or configurational relations have a distinguishing effect.

7.5.5.2 What about inversion?

I should say something about an influential approach to copular sentences. In a number of works Moro (1988; 1997) proposed a ‘unified theory of copular sentences’ based on what he sees as the syntactic property of inversion of the subject. In other words, all copular sentences are underlyingly predications; the difference is that at surface structure they are either inversions or not. In particular, specificational sentences are analysed as inverted predications. I don’t necessarily disagree with Moro’s analysis; after all, I just said in the immediately preceding paragraph that all copular sentences are syntactic predications. And furthermore, I accept that specificationals may be analysed as an inversion of the flanker that ends up in the pre-copula position (and there seems to be a consensus now that this is the case e.g. Heycock (2012)). And I also think that it’s fair to say that the mechanism that allows inversion is fundamentally syntactic in nature, an instance of Move, or internal Merge (see section 2.3.7). However, I don’t see how this is a theory of anything except the fact that a flanker has inverted or not. In particular, I don’t see how it makes the apparently meaningful distinctions that are required to distinguish the at least four classes of Higgins’s taxonomy. I’ll assume the burden is on the syntacticians to show why I might be wrong about this.¹⁴ Also, whilst I again accept that the fact of moving/internally merging a lexical item may be a consequence of the laws of syntax (in the same way as the fact of merging/phasally coinciding lexical items is a necessary component of syntactic predication), I’m not sure the theory of syntax alone tells us *why* inversion takes place on any given occasion. Other linguists, e.g. Birner (1994), have given general analyses of felicitous inversion; on one account inversion

¹⁴Authors on copular sentences often use the terms ‘specificational’ and ‘inverse copular sentence’ interchangeably e.g.: ‘Other approaches have reduced Higgins’s typology even more, positing only two classes: predicational and specificational (or inverse), depending on whether the more referential NP is the first or the second in the clause’ Arche et al. (forthcoming, 12). But they aren’t interchangeable; they relate to different ways of describing copular sentences. As we’ll see in the chapter 8 there can be non-inverse (‘reverse’) specificationals.

requires a preposed constituent to represent information ‘that is at least as familiar within the discourse as the postposed constituent’ (Birner *et al.*, 2007, 318). Such analyses relate to ‘information structure’, which we’ll look at in more detail later in this chapter, section 7.6. In the next chapter, 8.3.2, I’ll also look in some detail at what Hartmann (forthcoming) says about how inversion arises in specificational sentences.

Of course, it must be pointed out that dealing in the purely syntactic phenomenon of inversion may be precisely the right thing to do if one’s interests are in explaining phenomena that, superficially at least, seem to be purely syntactic. Much of the interest in copular sentences has been in explaining ‘connectivity effects’ (largely inspired by the work of Higgins himself) i.e. the way that we can invert (examples from den Dikken and O’Neill (2017)):¹⁵

- (15) a. What John_i is is important to himself_i.
 b. Important to himself_i is what John_i is.

but not

- (16) a. What John_i is is important to him_i.
 b. *Important to him_i is what John_i is.

Such effects were taken to be a diagnostic of a specificational (15) or predicational (16) sentence for Higgins. There are certainly straight syntactic reasons for this, but overall I would agree with Bejar *et al.* (forthcoming), who make the point that Higgins’s taxonomy, particularly the attempt to reduce it to two classes of predication and identity, is irrelevant in accounting for such syntactic phenomena. The fact that the connectivity effects correlate with differently meaning sentences does not explain why the connectivity effects occur.

7.5.5.3 What about word classes/phrasal categories?

What about word classes, lexical and functional phrases, like N, A, PP, NumP, and ClP? In chapter 1 we encountered Roy’s (2013) theory of predicational copular sentences. Roy’s theory is ultimately reducible to categorical differences, in particular whether the predicate is a NumP (maximal), a ClP (non-dense), or a non-nominal phrase—AP or PP—(dense). Whilst Roy wasn’t completely explicit about explaining the spatio-temporal profiles of predicates in syntactic terms, she did repeatedly emphasise the ‘very tight connection’ between syntax and semantics here, leading the reader to believe that a syntactic explanation was plausible.

¹⁵The subscripts indicate intended coreference. The unacceptability judgement of (16b) is, I take it, based on the intended coreference: *Important to him is what John is* is acceptable if *him* refers to someone other than John.

But let's go back to what we discovered in chapter 6. Roy's theory is essentially an account of the spatio-temporal profile of predicates, which in her terms is really the way we refer to an eventuality. This is as distinct to the way that we refer to an object, but there is a very neat parallelism between these ways of referring. As pointed out in section 6.5.1 the very same 'divisive' feature that accounts for mass noun reference when it applies to an object, produces dense predicates when it applies to eventualities (and is correlated with/explained by the presence of a NumP category in both cases). So, in sum, we can say that Roy's theory is a theory of ways we can refer to eventualities. So if I'm right about Roy's theory, and Roy herself wishes to at least correlate these different ways of referring with categories like NumP, CIP etc. and if Chomsky is right that referring is a *use* of language, see 6.3, then NumP, CIP etc. must be aspects of language *use* and not part of the unused UG or I-language of an individual.¹⁶ So I shall assume that the way of referring to a predicate (its spatio-temporal profile) may well be an important feature of the flanker of a copular sentence, but I'll consider it a use of I-language, rather than a primitive of syntax.

7.6 Information Structure

7.6.1 A need for something else

So far, we have encountered a number of properties of copular flankers that could be useful for making meaningful distinctions, and which I'll compile into a list below, section 7.7.2. There's one more area we need to look at though. Let's take a simple example:

(17) John is the culprit.

What kind of sentence is this? One thing we could do is analyse its flankers. Let's say *John* is a 'referentially used entity-denoting expression' and *the culprit* is an 'attributively used entity-denoting expression'. This is all well and good. But (17) seems to be ambiguous in a certain way. In terms of Higgins's sentence functions, it could either be predicational (it could be ascribing a property to John), or it could be specificational (it could be listing the culprits). To see this, notice the way that the sentence can felicitously respond to two

¹⁶I confess that I'm not completely comfortable with this conclusion. Intuitively, classifying, i.e. turning an object or eventuality into atomic parts, seems like something that relates to human cognitive psychological capacities. Having said that, we needn't necessarily lump this together with the capacity for hierarchical structure-building, which at least some authors wish to reserve the term 'syntactic' for, e.g. Boeckx (2014). Other grammatical theories bring in general features of cognitive psychology which may well be responsible for our capacity to 'atomise' objects and events, e.g. Langacker (1987). Note that this would make the use of I-language here an 'internal functionalism' in the sense of Hinzen (2012a) i.e. a use of one biological capacity in the mind by another.

different questions:

- (18) A: There's been a crime involving John. But is he the victim or the culprit?
B: John is the culprit.
- (19) A: There's been a crime. Do we know who the culprit is?
B: John is the culprit.

So we can see from this that we need to say something more about (17) in order to disambiguate it.

7.6.2 Back to Frege

We already encountered in our discussion of Frege's functions in section 2.2.3 a need for a notion of 'how to identify the function of a sentence'. There we saw that in a sentence like

- (20) Carbon dioxide is heavier than hydrogen.

there are potentially two functions: carbon dioxide is heavier than x , or x is heavier than hydrogen. The way to identify which one is in fact the function of the sentence is to know which function is the part of the expression that does not change, and which part of the expression is exchangeable. The only way to do this is to resort to an implicit comparison sentence or set of sentences. So if the comparison sentence is

- (21) Carbon dioxide is heavier than oxygen.

then the function is: carbon dioxide is heavier than x . But if it is

- (22) Helium is heavier than hydrogen.

the function is: x is heavier than hydrogen.

The problem of how to disambiguate (17) and how to identify the Fregean function of a sentence involves notions that linguists have elaborated under the rubric of 'information structure', which we'll look at some of the basic notions of now.

7.6.3 Basics of information structure

To start with, I'll just briefly set out the basic notions that enter into information structure. Afterwards, I'll say a few words about where I think this fits in with what I said about I-language, the use of I-language, and ostensive communication in chapter 6.

Information structure is fundamentally about the way in which we package information relative to a ‘common ground’ state of knowledge that is shared between interlocutors (Stalnaker, 1974). Common ground contents are the propositions and entities that are mutually known to be shared by interlocutors. There are three main ways in which this packaging is understood.

7.6.3.1 Focus

Focus relates to Rooth (1992)’s insight that any given linguistic expression is related to a set of alternative linguistic expressions. For example:

(23) John ate the cake.

The alternatives here could be that someone else ate the cake, that John did something other than eating to the cake, or that John ate something other than the cake. This follows from the way that information generally ‘reduces uncertainty’ in Shannon (1948)’s terms. If one knows that someone has eaten the cake, but one doesn’t know who did it, then the set of possible alternatives to *John* eating the cake (Bill eating the cake, Mike eating the cake etc.) is an uncertain state of knowledge. Or it is a less certain state of knowledge than the state one occupies when one subsequently is told (23). In this context, then, we say that the focus is on *John*—John is the selected alternative from a set of alternatives—and that the ‘background’ is the set of alternatives to John. This can also be understood in terms of the question under discussion (Benz & Jasinskaja, 2017). Propositions can be quite generally understood in this way. Given the context we just gave, the question under discussion for (23) is:

(24) Who ate the cake?

The answer to the question selects an item from a set of alternatives, which is the focus of the answer.

7.6.3.2 Givenness

Féry and Krifka (2008, 4) define givenness as follows:

A feature X of an expression α is a Givenness feature iff X indicates whether the denotation of α is present in the common ground or not, and/or indicates the degree to which it is present in the immediate common ground.

Information which is not given is ‘new’. For von Stechow (1999), the given–new distinction is a reflex of a more general contrast in informativeness that can apply at the level

of the sentence or the level of the discourse context of the sentence. Given information (the less informative part) = *already known*; new information (the more informative part) = *not known* or *newly introduced*. This is more of a scale than a binary distinction. Whilst there is certainly a great deal of overlap between new information and focused information, they are distinct categories. The evidence for this is that pronouns, which are always necessarily given (they relate to antecedents that are present in the common ground), can be focused (example from Krifka and Musan (2012, 25)).

(25) Mary only saw HIM.

7.6.3.3 Topichood

Reinhart (1982) makes the point that the common ground is not an unstructured set of propositions and entities. The common ground is rather like a file card storage system, where the file cards have headings (which are entities or sets of entities) which the common ground contents are stored under. As such, when one enters a proposition into the common ground (i.e. when one adds new information into the common ground), one does so with a heading, and it is this heading entity/set of entities that is the topic of a sentence. This is intended to chime with the way that human memory works; propositions are organised in memory under headings, though it isn't necessary that information is stored in this way ('...relational databases or sets of possible worlds, both models of information, do not presuppose any relation of aboutness' (Krifka & Musan, 2012, 27)). In this way, two propositions can express the same information, but should be stored differently:

- (26) a. [Aristotle Onassis]_{Topic} [married Jacqueline Kennedy]_{Comment}.
 b. [Jacqueline Kennedy]_{Topic} [married Aristotle Onassis]_{Comment}.

(26a) should be stored under the 'Aristotle Onassis' heading, and (26b) under the 'Jacqueline Kennedy' heading. Topics are often given information, inferable from discourse context, but they can be new information too. Krifka and Musan (2012, 28) give the following example, saying that it 'introduces a new entity into discourse and, at the same time, uses it as the denotation of a topic constituent, which amounts to introducing a new file card in the common ground content':

- (27) [A good friend of mine]_{Topic} [married Britney Spears last year]_{Comment}.

And whilst the comment and focus are often the same, they need not be identical. In the example below, the comment contains the focus:

- (28) A: Tell me something about Onassis. When did he marry Jacqueline Kennedy?
 B: [He]_{Topic} [married her [in 1968]_{Focus}]]_{Comment}

For von Heusinger, topics are the sentence-level reflex of the more general contrast in informativeness that applies to sentences themselves as well as the level of the discourse context of sentences. The topic is the less informative part and the comment the more informative part (von Heusinger, 1999, 102).

7.6.3.4 What is information structure?

It isn't entirely obvious to me where 'information structure' fits with I-language, the use of I-language, and ostensive communication (see chapter 6). With respect to the basic division between more or less informative parts of a sentence that von Heusinger considers to be at the root of information structure, he says: '[T]here is no agreement where to locate this level. For some researchers, this dichotomy belongs to pragmatics and text linguistics, for others it is part of the psychological states of the participants in a conversation, and others integrate it into grammar proper' (von Heusinger, 1999, 102). Scott-Phillips (2015b, 68) points out that the mindreading (theory of mind) skills which make ostensive communication possible are also required to establish and successfully contribute to the 'common ground'. We must provide appropriate information for a particular audience which means we must be able to represent other minds. This makes me think that information structure could be described as a use of the cognitive capacities that give rise to ostensive communication, moreso than as a use of I-language directly. This would predict that we should be able to come up with examples of information packaging in the absence of I-language, though I confess I can't think of any. If Reinhart's file card metaphor really does reflect how human memory works, then the fact that sentences have topics may be related to this. Also, as Féry and Krifka (2008, 12) point out, '...the tendency of topics to be prosodically separated from and to precede the rest, can be traced back to the optimal flow of information transmission'. Finally, as mentioned above, the fact of selecting an item (the focus) from a set of alternatives appears to relate to the general notion of uncertainty reduction in information theory. So my conclusion is that information structure may relate to a mixed bag of psychological and non-psychological phenomena. As such, I think it's safest to categorise it separately from any of the other 'pieces of the puzzle' collated below in 7.7.2.

7.7 Multiple taxonomies and pieces of the puzzle

7.7.1 Multiple taxonomies and a ‘full description’

By now, we have inspected a number of different areas in order to see if they can shed light on a description of the taxonomy of copular sentences. The first thing to say is that, generally, taxonomies are theorists’ groupings of phenomena according to their interests. So there will be different taxonomies of copular sentences according to what properties we’re interested in about them.¹⁷ Most of the properties we’ve encountered so far are meaningful in some way, but the inversion vs. non-inversion property in and of itself isn’t doesn’t appear to be. And yet, if that is what one interested in, then copular sentences can indeed be grouped into two categories in this way. If one’s interest is in metaphysical categories, then this is a way one can potentially describe at least some of the copular sentences; if it’s the spatio-temporal properties of nominal flankers (that enters into the i-level/s-level distinction), you can divide them up as Roy (2013) does. It’s even possible that one might want to relate some of the copular sentences to representations in predicate logic, though the work of chapters 2 and 3 attempted to show that this should be of little explanatory interest.

Having said this, I think that a ‘full description’ of copular sentences will be the conjunction of all the properties that enter into determining distinctions in meaning between them. Let’s just look at a toy model with two properties A and B that can apply positively or negatively to each flanker. Let’s mark this with a superscript $A+$, $A-$, $B+$, or $B-$ next to each flanker. Furthermore, let’s mark the pre-copula flanker as ‘Flanker1’ and the post-copula flanker as ‘Flanker2’. In this situation the number of logically possible permutations is 16, the first few lines of which would be:

1. Flanker1^{A+B+} *is* Flanker2^{A+B+}
2. Flanker1^{A+B+} *is* Flanker2^{A+B-}
3. Flanker1^{A+B+} *is* Flanker2^{A-B+}

The total number of permutations can be calculated simply using the rule of product: There are two ways in which A can be valued on Flanker1 (+ or -) and two ways in which B can be valued on Flanker1 (+ or -). $2 \times 2 = 4$. So there are four ways that Flanker1 can be

¹⁷This principle of multiple taxonomies can be seen in recent work about ‘how the different types of copular structures are differentiated’, which can be, ‘either lexically—that is, through the use of distinct copulas—syntactically through different base structures or different derivational sequences, or semantically, deriving the different interpretations from the contributions of the subjects and predicates’ (Arche *et al.*, forthcoming). Here, though, I have argued that there is little hope for the ‘lexical’ differentiation of the copula or a syntactic differentiation of base structures.

valued. Flanker1 and Flanker2 are identical with respect to possible permutations, so there are also four ways that Flanker2 can be valued and therefore there are $4 \times 4 = 16$ possible permutations of Flanker1 and Flanker2 together.

This is how I think we should think about any putative copular sentence taxonomy. Each permutation corresponds to what is effectively a different type of copular sentence, and in principle could deliver a different reading. What is immediately striking is that even with just two positively or negatively valued properties on each flanker, the number of potential readings—16—is considerably more than the four (Higgins) to six (Roy) types of copular sentence that are acknowledged in the literature. Showing that a larger number of different readings is in fact available is evidence for this approach and something I aim to at least sketch out by the end of this thesis. So summing up the approach here, we can say that it leads to a more pluralistic description than the four to six types of copular sentence that are usually acknowledged. I've aimed to show that the burden of explanation is on those who wish to reduce the taxonomy to a smaller number of superordinate categories to show the value of doing this.

So the question then becomes: what are the flanker properties that I resort to? What are the real equivalents of A_{\pm} and B_{\pm} of the toy model?

7.7.2 Pieces of the puzzle

The following list compiles the factors that we have encountered so far. With respect to groupings, I confess that there are conceptual difficulties. We've just encountered the information structure properties, whose status was unclear, so they'll be grouped separately. I won't bring metaphysical properties into the case study in the next chapter because the case I look at involves a pair which I'm assuming are the same metaphysically. However, I'll confess that I'm unclear about how I would view the descriptive role of metaphysical categories. Let's take two cases of real properties, in the sense of section 7.3:

- (29) a. Dobbin is a horse.
 b. Dobbin is brown.

(29a) is an instantiation of a kind in Lowe's (2009) terms, and (29b) is an attribute exemplification. Now clearly metaphysical categories and the relations between them are facts about mind-external reality. However, in terms of grouping the factors below, I think they count as a use of I-language in order to refer to things in the world. They are *what* we refer to, as opposed to *how* we refer.

Factors that I believe enter into either the description of copular sentences are marked

with a tick. Factors which enter into the discrimination of meanings are marked with a double tick.

- Syntax
 - Syntactic predication (reducible to phasal coincidence, section 7.5?) ✓
 - Inversion vs. non-inversion of flankers from original underlying syntactic position
 - Small clause restructuring giving rise to distinct scope relations between quantifier phrases (see section 5.2) ✓✓
- Predicate logic: predication vs. identity (chapters 2 and 3)
- Metaphysics: numerical identity (section 3.2)
- Use of I-language: words, sentences, referring (chapter 6)
 - Sentences: instances of manipulating an audience’s mental representations for extended social navigation (section 6.4.1) ✓
 - Referential uses
 - * Spatio-temporal profile of an eventuality i.e. referring maximally/densely/non-densely i.e. modifying an event argument with a NumP, ClP, or non-nominal (Roy, 2013). ✓✓
 - * Referential or attributive nominal reference (Donnellan, 1966) ✓✓
 - * Reference to the metaphysical categories of the flankers (e.g. objects, properties, kinds, tropes, see section 7.3; also eventualities/Davidsonian event arguments in at least some copular sentences, see chapter 1; also ‘senses’ of objects in Frege’s ‘On Sense and Reference’, see section 3.3.2, and possible worlds for intensional reference generally, see 2.2.6) ✓✓
 - * Other ways of referring (correlating descriptively with different phrasal categories) e.g. definite/indefinite (von Heusinger, 2013). ✓✓
- Information structure (section 7.6)
 - Focus and background ✓✓
 - Given, accessible, and new information ✓✓✓
 - Topic and comment ✓✓
 - The *reason* for flanker inversion (Birner, 1994) ✓✓

7.8 Conclusion

This chapter has tried to collate the conclusions that we've drawn so far in the thesis and deal with issues that were outstanding from previous chapters, as well as supplement the conclusions with some additional factors that enter into a full description of copular sentences that distinguishes their meanings. The overall picture can be viewed in two ways: (1) we can say that there are multiple taxonomies of copular sentences that correspond to our interests, be they in the metaphysical categories of the flankers, in the inversion of the flankers, in the information structure of the flankers etc. (2) we can say that there is a 'full description' of the properties that enter into making meaningful distinctions along the lines of the flanker-based taxonomy model that I sketched out in 7.7.1. The properties that are marked with a double tick in 7.7.2 can essentially apply to the flankers or not, and the taxonomy of copular sentences will be as pluralistic as the different number of sanctioned permutations of these properties. I'll exemplify what I mean by this in the next chapter. This approach is, I think, consistent with the general scientific endeavour of 'multi-factorial modelling', which in linguistics has recently been applied to pronoun form selection by Kibrik (2011). The basic idea is that complex phenomena are often the result of simple interacting parts, rather than relying on 'black box' models that usually only restate the phenomena to be explained/described. Disease causation, for example, is generally now viewed as a probabilistic abstraction over multiple interacting variables, rather than the result of any single deterministic cause (Parascandola, 2011). The comparison isn't exact: the multifactorial model in the domain of disease causation aggregates factors in order to predict the likelihood of disease. Here, the multifactorial model simply results in copular sentences that can be described as the aggregate of all the factors that enter into the description of their flankers.

The next chapter, then, is a case study that takes this basic picture and applies it to a small set of data in order to see how it works in practice.

Chapter 8

Putting it all together: a case study

8.1 Introduction

This thesis is not a work of first-order linguistics. The point, rather, has been to try to analyse some of the assumptions that are held by those who work on the problem of explaining or describing the meanings of natural language sentences. As such, the focus has, partly, been on trying to clarify the conceptual landscape which forms the background assumptions of explanations and descriptions of natural language sentences. There are many areas where there are still conceptual difficulties, so that giving a detailed analysis of all the different types of copular sentence is not a task that I would attempt at this stage. However, I think it's time to pin my colours to the mast and show how I would go about doing such an analysis. Rather like chapter 5, then, I'll focus on a case study to serve as a proof of concept, of how the approach to describing the copular taxonomy advocated in chapter 7 can be implemented. To do this, I'll just focus on a pair of English sentences:

(1) The culprit is John.

(2) John is the culprit.

Sentences like (1) are generally considered specificational copular sentences, where (2) is predicational (Mikkelsen, 2011).¹ (3) would be a typical context of utterance:

(3) At the scene of a crime:

Detective: Sargeant, do we know who did it?

Sargeant: Yes detective; the culprit is John Nettles.

¹In studies that are concerned with more purely syntactic properties of copular sentences, sentences like (1) are simply called 'inverse copular sentences' (Moro, 1997; den Dikken, 2006a), see section 7.5.5.2.

Now, a question immediately arises about the status of (2). Is this not basically the same as (1)? Clearly they are different—they have different word orders—but does this difference (in word order) mean that (2) is *not* specificational in whatever way it is that (1) is? That is the question I want to try and answer here, and in so doing I want to demonstrate the way that factors from chapter 7 can be brought in to analyse copular sentences. The answer will be that (2) sometimes is specificational, and so the word order difference doesn't necessarily discriminate them in that respect. But there is still a difference between them with respect to the way that (1) is necessarily a specificational with an exhaustive reading of the focused constituent of the sentence whereas (2) allows for a non-exhaustive reading (as well as the exhaustive one). The essential difference between 'specificational' and 'predicational' will be understood as a result of a combination of information structure and the *way* we can refer, which means that we can conclude that *John is the culprit* and *The culprit is John* may be semantically and syntactically identical.² Certainly, there will be no need to think of the difference between (1) and (2) as relating to the logical difference between predication and identity (which will be consistent with the way I've tried to show in this thesis that this is not generally relevant to analysing copular sentences).

Before moving on, one side-issue is worth remarking on. There is also a potential equative reading for both (1) and (2). Wiggins (1965, 43) recognises this as possible in an 'abnormal context'. For the sentence

(4) Plato was the author of the *Sophist*.

Wiggins says, '... it might in some very special context be intended to express the genuine identity-statement that a definite logician, the author of the *Sophist* (an individual independently identified), and Plato (independently identified again) were one and the same person'. The intuitive way to see this is simply to think of '*the author of the Sophist*' (or *the culprit* in (1) and (2)) as functioning like proper names, in the absence of one actually knowing the proper name for a person one has 'independently identified' or that one has 'acquaintance' with, in the sense of Russell (1912, 46–60). Barros (2016, 2) confirms this reading: 'In [*Jack is the president*], [*the president*] may be interpreted as referential or predicative. On the predicative reading, [the sentence] counts as a predicational clause. . . on [*the president's*] referential reading, [the sentence] counts as an equative clause'. This all just shows that there is a potential equative reading of (1) and (2) and we don't want to conflate them with the 'predicational' and 'specificational' readings that we'll be dealing with in what follows.

A starting point view on this issue would be this: (1) is specificational, not predica-

²Modulo the word order difference; see below, section 8.4.3, for a suggestion about *why* this difference may be necessary.

tional, and (2) is predicational, not specificational. The debate about specificationals (see Mikkelsen (2005)) includes this possibility but has mainly been about whether they can ultimately be defined as inverted predicational sentences or equative sentences (Moro, 1997; Heycock & Kroch, 1999). Most authors agree that (2) is predicational, or rather, can have a predicational reading as one of its readings. Some, though, disagree; for example, Roy (2013, 12) says: ‘Cases involving a definite DP in post-copular position, and which generally allow inversion. . . fall, in my view, into the class of non-predicational sentences’.³

At this point, it’s absolutely crucial to say that the literature recognises ‘reverse specificational’ sentences in order to refer to specificational versions of sentences like (2) (see Patten (2016, 4)). So (2) is quite generally accepted to be ambiguous between a predicational and specificational reading (though I think this fact is sometimes overlooked). So in a way, the question of this chapter could be: In what way is (2) ambiguous such that it can be both specificational and predicational?

In the rest of this chapter I’ll firstly outline where I think the specificational vs. predicational difference comes from, 8.2—essentially it is the fact of which flanker is focused. This allows us to hold on to the view that (1) and (2) are identical in their syntax–semantics. However, there is a potential counter-example, 8.3, sentences which seem to show that there is an apparently semantic distinction between (1) and (2) even when they are both specificational (so where (2) is a ‘reverse specificational’). I look at two previous analyses of the source of this apparently semantic fact in 8.3.1 and 8.3.2. 8.3.2 also gives the background to Hartmann (forthcoming)’s analysis of the difference between specificational and predicational sentences quite generally. I spend some time on this because I think it is one of the most comprehensive recent analyses of copular sentences which is in the style of the sort of analysis that I think will be successful, and is in line with how I have argued copular sentences should be analysed in this thesis. However there are some issues with it which prompt me to offer an alternative analysis in 8.4. At the end, section 8.4.6, I return to Higgins’s eight-ways ambiguous sentence which I presented in the introduction, section 1.3, to see how my system fares in the light of it. In fact it raises complications for the system presented here which demonstrate that further work is required.

³Later on however, (Roy, 2013, 59), Roy appears to contradict herself when she says: ‘In (75) [*Je me croyais le meilleur*—gloss: I myself believed the best—trans: ‘I believed myself to be the best one.’], the small-clause predicate is a definite expression, used predicatively, showing that the predicate position inside the small clause is not restricted to bare expressions only’. She may be making a relevant distinction between the terms ‘predicational’ and ‘predicative’ (though as far as I’m aware this is merely a difference whereby ‘predicative’ is used to describe a type of sentential element which is a predicate—i.e. a predicative adjective *blue* in *My car is blue*—and ‘predicational’ is used to designate *sentences* which are defined by the presence of such elements).

8.2 The source of two readings—predicational and specificational

Patten (2016) is unusually explicit on the question of the relationship between (1) and (2). In order to give some flesh to the bones of what ‘predicational’ and ‘specificational’ stand for, let’s go for an initial rough characterisation along lines Patten herself gives.

- (5) A **predicational** sentence ascribes a property to a referent (Patten, 2016, 2).
 A **specificational** sentence functions like a list, and says what item(s) fall under a given list heading (Higgins, 1979); specificational sentences involve a ‘class inclusion’ relation which, ‘serves to specify the membership of a category’ (Patten, 2016, 2).

This is very much in line with the mainstream characterisation of (1) and (2). Interestingly, Patten (2012, 56–7) says that these are ‘opposite perspectives’, adding ‘specificational meaning... is quite literally the *inverse* of predication’. Patten’s own proposal (following Croft (1991)) is that there is a ‘nominal predication relation of class inclusion’ that is common to both predicationals (with nominal predicates) and specificationals.

So in the predicational *John is the culprit*, the sentence **says what class John is included in**. In the specificational *The culprit is John*, the sentence **says what is included in the class of culprits**. An important point to note here, though one which I won’t pursue, is that the definiteness of the set designated by *the culprit* does indeed make an important difference, say between the sentence *John is the culprit* vs. *John is a culprit*—which could similarly be said to be saying that John is included in the class of culprits. So what one really means here for the version with the definite article is that John is a member of a contextually salient set of culprits (Patten, 2016, 5); for the example of *The waitresses are Diane and Carla*, Patten notes that they would have to be waitresses in a particular bar, say, that would be contextually salient to speaker and hearer. As well as contextual salience, Patten associates definiteness with an ‘inclusive’ reading, whereby *the culprit* or *the waitresses* refer necessarily to *all* entities in the contextually salient set, so *all* the waitresses in a particular bar or *all* the culprits of a particular crime (which if it is singular *the culprit* will only be one individual).

8.2.1 Focus

Patten’s analysis can be understood in terms of the notion of focus that we encountered in section 7.6.3.1. For the predicational reading, the focus is on *the culprit*. The background set of alternatives is all the things John could be (the culprit, the victim, a bystander etc.);

putting the focus on *the culprit* selects this item from the set of alternatives. Alternatively, the focus can be on *John*. Here the background set of alternatives are other people that could be the culprit (Bill, Mary, Keith etc.); putting the focus on *John* selects him from the set of alternatives. Another way to put this is in terms of ‘questions under discussion’ (QUDs), cf. (Milway, 2016). In the case of the predicational reading, the question under discussion is, say, ‘What class is John included in?’, or (abstracting from *John*) ‘For a given $x_1 \dots x_n$, what is the P, such that $x_1 \dots x_n \in P$?’; for the specificational on the other hand, the question under discussion is ‘Who or what are included in the class of culprits?’, ‘For a given P, what are the $x_1 \dots x_n$ such that $x_1 \dots x_n \in P$?’.

So it seems that in terms of a standard formal semantic account of the meanings of *John is the culprit* and *The culprit is John*, we can indeed say they are the same. Put another way, we could say that the ‘class inclusion’ relation whereby John is a member of the class ‘the culprit’ is the same in both cases. The difference between the predicational and specificational reading resides essentially in the difference between what is presupposed in the common ground of the interlocutors. If we presuppose that John is a member of a class, but wish to inform someone about which class it is, then the sentence we produce is predicational. If, on the other hand, we presuppose that there is a particular class of culprits, but wish to inform someone about the membership of this class, then the sentence we produce is specificational.

Do we have to produce the words in a different order to signify this? i.e. do we *have* to say *The culprit is John* for the specificational reading and *John is the culprit* for the predicational reading? Or can we also say *John is the culprit* for the specificational reading? I will accept the standard view that the (1) order can never be predicational as a stipulation for now, so the question of a potential ambiguity only arises for the (2) order. A’s question in (6), below, sets up a specificational context (capitalisation indicates nuclear tone):

- (6) A: There’s been a crime. Who is the culprit?
 B: The culprit is JOHN
 B’: JOHN is the culprit.

B’s response in (6) certainly seems like a perfectly felicitous reply (as well as B’s response, with the classic ordering of a specificational). In fact it isn’t obvious why we would choose B over B’ in (6), though the necessary exhaustivity of the B/(1) ordering—which we go into some detail about below—is at least part of an answer to this question. As I say, Patten (2016, 4) is explicit about the position she takes here. She says

If the subject is placed in focus (marked by intonation in (11)), the sentence acquires a specificational reading: the speaker is listing the membership of the

category of waitresses.

(11) DIANE and CARLA were the waitresses.

and she notes that examples like her example (11) are, as we noted earlier, sometimes called ‘reverse specificational sentences’, ‘... since they share the same specifying function as sentences like (7) above [*The waitresses were Diane and Carla*], but not the characteristic inverse word order’ (Patten, 2016, 4).

8.2.2 Referential vs. attributive

So, we have seen a way of distinguishing the two readings of (1) and (2). Both readings are available in (2). Only the specificational reading is available in (1). Now, I don’t actually think it is necessary to state a specific ‘semantic relation’ between the flankers. Partly because I don’t know what semantic relations are generally, but also because I think that the phenomenon that is at stake here can be described in terms of a property of the flanker. The crucial difference between *the culprit* and *John* (and also between *the culprit* in (1) and (2) and *the culprit* in an equative version of (1) and (2)—recall the possibility of this ‘abnormal’ reading we gave at the end of section 8.1) is in terms of *the way that they refer*. Donnellan (1966) made a distinction between ‘attributively’ referring definite descriptions and ‘referentially’ referring ones (Higgins (1979) also thinks the distinction applies to proper names, and I’ll follow him in this). The attributive use is used where we assume that there must be **someone or other** who committed a crime, but we don’t have a particular person in mind. Higgins referred to the distinction as the difference between using a noun phrase with ‘acquaintance’ or not, in a sense taken from Russell (1912; 1917). We often don’t know what particular person or thing is picked out by a definite description, but it can be useful to refer to someone or something, *who/whatever* it is, on the basis of a description. The simple example that we’ve used throughout is a paradigm case: in the setting of having discovered a crime, of being at the scene of a crime, we can plausibly deduce that there is a culprit because crimes are always committed by culprits. So we can refer to this hypothetical entity even though we’ve never met them or know anything about them. The evidence of criminality (the fact, say, that some items have been stolen from someone’s house, their door broken down etc.) tells us that there is a culprit somewhere in the world. We can think of *the culprit* as referring to ‘putative’ objects, objects that must exist but which we aren’t in fact acquainted with. If we are so acquainted, then we can use a definite description to refer ‘referentially’, rather as a proxy for a proper name.

So, instead of talking about a ‘class inclusion’ relationship between *the culprit* and *John*, I’ll simply talk about *the culprit* being attributively used, and *John* being referentially used.

In other words, I'm using Donnellan's distinction between an attributive and referential flanker instead of Patten's (due to Croft) distinction between a class and a class member. I confess that it is more intuitive to talk of relations, but I think the referential–attributive distinction is better-grounded in the philosophical literature and it does the work that the class inclusion relation is brought in to do.

So we've made a basic distinction between a predicational and specificational reading, which can be delineated in terms of the ways the flankers refer (*the culprit* = attributive, *John* = referential) and which flanker is focused (or the distinct QUDs they address). But these don't really seem to make a semantic distinction between (1) and (2), properly speaking. As I said in chapter 7, they are to do with the referential *use* of I-language, and information structure. In that case, therefore, we can stick to the line about these sentences that they are still essentially equivalent semantically. In the next section, we'll deal with a counter-example to this position.

8.3 A counter-example to the view that (1) and (2) are semantically equivalent

From Patten's (2012; 2016) account, then, it seems that (1) and (2) can have equivalent readings, and the difference between the predicational and specificational reading is a matter of information structure, not their semantics properly speaking ('perspective' differences as Patten puts it). This section will focus on an apparent counter-example to this view, involving possessive constituents and the additive particle *too*. I should say immediately that even though we are dealing with possessive subjects like *John's friend* or *Mary's brother* in the counter-examples below, I am just assuming that sentences like *John's friend is Susan* have a specificational reading in the same way as (1), *The culprit is John*. We just don't see the sort of exhaustivity phenomenon that we shall see in the examples just below when we use definite descriptions like *the culprit*. There is surely a reason for this, to do with the nature of definite descriptions in English (i.e. they basically just seem to presuppose the uniqueness of their referent, as has been recognised since Russell (1905): 'The *F* is *G*' says that there is an *F* and at most one thing is *F*). But the point of the section is to show that there are examples which appear on the face of it to reflect proper semantic differences between the (1) and (2) orderings, notably differences that putatively arise due to one being a logical predication and the other being a logical identity. However as we'll see below, the examples really all seem to be instances of specificational readings (in Patten's sense)—i.e. one is a classic specificational ((1) ordering) and the other a 'reverse specificational' ((2) ordering) in

standard terms. The difference between them relates to an exhaustivity requirement that, again, I wouldn't think of as a properly semantic distinction moreso than an additional information structure parameter.

8.3.1 A 'well-noted maximality difference' with pre- vs. post-copula possessives

Selvanathan (2016, 241)—following Partee (1986b)—points out a 'well-noted maximality difference' with respect to possessives, depending on whether they are used as 'arguments' (Selvanathan's term) or 'predicates'; this can be seen in the examples below:

- (7) a. Susan is John's friend, and Mary is John's friend too.
 b. */?John's friend is Susan and John's friend is Mary too.

(7a,b) both try to express the fact that there is another person who can also be described as 'John's friend'.⁴ The relevant contrast between (7a), on the one hand, vs. (7b) on the other is this: in (7a) *too* can value the variable in the QUD 'What is the x where x is a friend of John?'; the second conjunct of (7b) cannot do this.

Selvanathan's account of the difference between acceptable (7a) and unacceptable (7b) is to say that possessives are semantically ambiguous between two interpretations, represented as:

- (8) *John's friend*
 a. λx friend-of (x , John)
 b. ιx . friend-of (x , John)

(8a) is a standard predicate—an $\langle e,t \rangle$ function—which will be satisfied by taking an argument to give a true or false expression. (8b), on the other hand, is an object (an $\langle e \rangle$ -type expression); it is 'the unique object x such that x is a friend of John'. Now, in order for an object like (8b) to function as a predicate, it needs to be type-shifted into an $\langle e,t \rangle$ format using the IDENT type-shifting operation from Partee (1986b, 209), which creates an

⁴There is a reading of the clause *Mary is John's friend too* and a reading of the clause *John's friend is Mary too*, neither of which is relevant here, and which are disambiguated by the context anyway. *Mary is John's friend too* could express that John is another person who Mary is friends with, which would be possible in a context where we were talking about people who Mary is friends with i.e. *Mary is JAMES's friend and Mary is JOHN's friend too*. Note that in this context placing the *too* directly after *Mary* is less felicitous: *?Mary is JAMES's friend and Mary, too, is JOHN's friend*. *John's friend is Mary too* might also be felicitously uttered in a context where we are talking about people whose only friend is Mary (QUD 'What is the x where x 's only friend is Mary?') e.g. *BILL's friend is Mary and JOHN's friend is Mary too*. This, again, is not the reading we are interested in, and (7b) doesn't allow for it anyway, starting as it does with *John's friend* and not another name that could be contrastive with *John's friend*.

equation, shifting (8b) from: ιx . friend-of (x , John) to: $\lambda x [x = \iota y$. friend-of (y , John)].⁵

In (7b) it is the type-shifted semantic representation of (8b) that is used for *John's friend*, so that the first conjunct of (7b) essentially reads, ‘the unique x such that x is the friend of John = Susan’. From here it is clearly going to be a contradiction to introduce another unique friend of John, and it is thus that the oddness of (7b) arises for Selvanathan. Why must it be that (8b) is the representation of (7b)? Well, this is just the way Selvanathan’s analysis works. Essentially, for him the ultimate stipulation is that (7b) is specificational and specificationals are equations. As such, (7b) only allows for a semantics of equation, so (8b) must be the representation of *John's friend* in (7b). It would not be possible to achieve a semantics of equation by using (8a). Combining (8a) with an object, say Susan, would ultimately give us ‘friend-of (Susan, John)’ which is not an equation; it is just a straight predication, and we can easily see how it wouldn’t give us the kind of uniqueness meaning that makes (7b) odd. ‘friend-of (s , j)’ just translates into English as ‘Susan is a friend of John’, and there is clearly no implication that Susan is the only friend of John.⁶ The goodness of (7a) is simply that we are naturally assuming that the (8a) denotation is the intended one for *John's friend* (as it is available and is the only one that makes sense); we just can’t do this for (7b).

It’s important to highlight that neither of (7a) or (7b) are predicational in the terms we established in section 8.2. There, following Patten, we said that ‘specificational’ essentially corresponded with ‘focus on the referential term’ and ‘predicational’ with ‘focus on the attributive term’. But here focus is on the referential term in both cases. We are therefore dealing with an instance of a specificational (7b) and a ‘reverse specificational’ (7a), not an instance of what would standardly be considered a predicational copular sentence and a specificational one (as Selvanathan seems to think). But there is an apparently semantic difference arising from the word order difference between (7a) and (7b), and this needs to be explained if we are to hold on to the view that pairs like (1) and (2) are essentially semantically equivalent modulo information structural differences.

8.3.2 Hartmann (forthcoming)

A different analysis of such sentences as those in (7) is given by Hartmann (forthcoming), who follows Heycock and Kroch (1999) in considering specificationals to be inversion structures

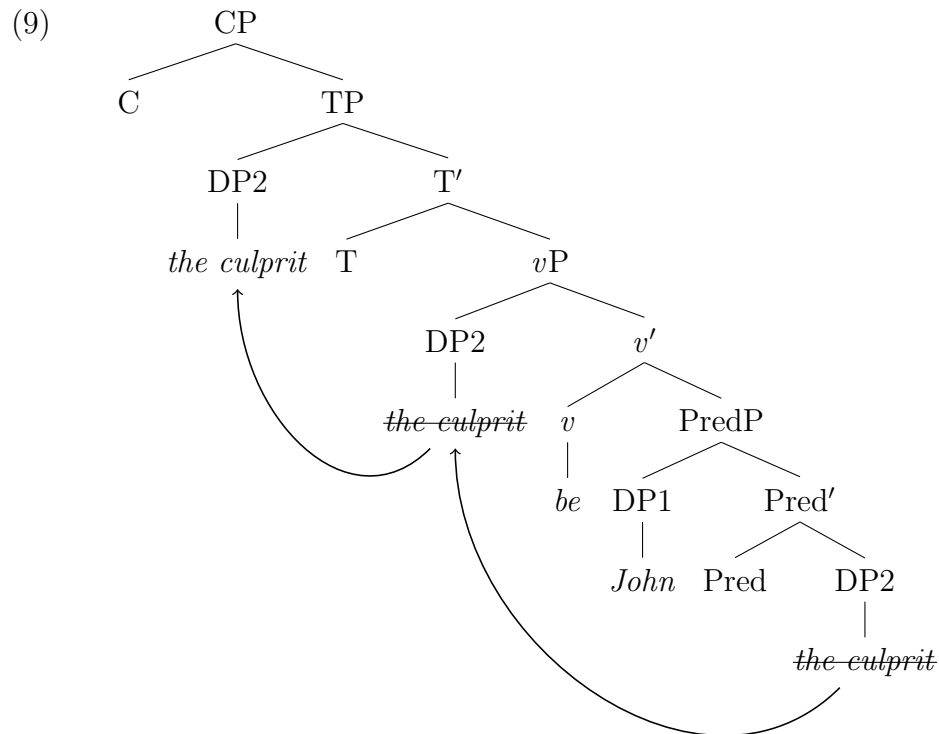
⁵See section 2.3.6 for the notion of type-shifting.

⁶Selvanathan doesn’t ban the possibility of the object/equative reading for the *Susan is John's friend* order; he is quite explicit that we should regard this as ambiguous (‘When the possessive occurs as a predicate, both denotations are possible’ Selvanathan (2016, 241)). It’s just that the *John's friend is Susan* order is not so ambiguous.

where what is inverted is not a predicative ($\langle e,t \rangle$) expression as such, but rather a DP (so an $\langle e \rangle$ -type object expression of some kind).

But let's start with her analysis of the basic cases in (1) and (2). For *The culprit is John* and its specificational reading, Hartmann says the underlying subject (*John*) 'stays low', meaning that it does not rise into the Spec,TP position—the canonical subject position (which I previously referred to as Spec,IP in chapter 2, see footnote 11); the reason for this, and for why the inversion of *the culprit* takes place at all is what Hartmann calls 'Focus-Background Mapping'. Focus-Background Mapping occurs 'in the core predication', which is to say at the level of a Predicate Phrase (PredP) that is the nucleus for the rest of the derivation. The specifier of PredP is mapped to the focus and the complement of PredP is mapped to the background of the sentence. Hartmann says (16): 'I take this mapping to be triggered by a type mismatch of two definite and referential DPs'. Two $\langle e \rangle$ -type DP flankers represent a type mismatch that must be resolved; I think the logic of Hartmann's argument here is that this is because functional application needs one flanker to be an argument of the other. She says (18): 'This type mismatch can be resolved by assigning a functional meaning to the complement of [Pred]. This functional meaning is what the background expresses'. It isn't clear to me whether this is the same thing as a kind of semantic type-shifting (see 2.3.6). Hartmann goes on to note that 'true equatives' also seem to display a functional interpretation of one of the flankers (as the comment of a topic–comment information structure).⁷ But I confess that what the nature of the 'type mismatch' is and why it should trigger this functional mapping/interpretation is not entirely clear. In any case this mapping, in English, assigns a formal feature to the background constituent i.e. the complement of PredP here, which feature drives the inversion of the background/complement of PredP to Spec,TP (via Spec,*v*P) to become the superficial subject of the sentence. The syntax of specificational copular sentences Hartmann (forthcoming, 17) gives is:

⁷ Hartmann and Hegedűs (2009) also make the point that in Hungarian equatives 'do not show a neutral order: they either exhibit a topic–comment, or a focus–background structure' (Hartmann, forthcoming, 18). Interestingly, these facts along with Hartmann's point about type mismatches may be an argument against representing equatives with the '=' sign of predicate logic, as it would seem that syntax cannot tolerate copular sentences with two $\langle e \rangle$ -type flankers without making one of them a different type somehow (albeit a different information structure status rather than a semantic type necessarily; as I say, it isn't clear to me what relation Hartmann intends there to be between semantic types and information structure). Equations with the '=' sign, on the other hand, surely do equate identically typed flankers.



Where *John is the culprit*—a ‘predicative copular clause’ for Hartmann—is concerned, Hartmann’s analysis is somewhat different. But the difference is not that *John* is not focused in these examples. One difference is that *the culprit* is not a DP (it is rather a NumP). The Focus-Background Mapping that is triggered by the presence of a ‘type mismatch’ when there are two DPs does not apply. In *John is the culprit* therefore, *John* can be assigned focus but *the culprit* is apparently not assigned as the background. There is also a ‘discourse’ level difference: in *John is the culprit*, *the culprit* is marked as ‘given’ at the ‘discourse’ level, as is *John* (in the context of the question *Do you think John or Bill is the culprit?* Hartmann (forthcoming, 18)).⁸ In Hartmann’s specificational sentence case, *the culprit* is marked as ‘accessible’ (in the context of the question *What’s new in the murder case?* Hartmann (forthcoming, 15)) and *John* is marked as ‘new’. To see these differences more clearly, see Tables 8.1 and 8.2. The syntax of Hartmann’s ‘predicational copular clauses’ (Hartmann, forthcoming, 19), then, is given as (10):

⁸As we saw in the chapter 7, the given vs. new distinction is generally categorised as a fact about information structure, along with focus, so it isn’t entirely clear whether ‘discourse’ is something distinct from information structure in Hartmann’s analysis.

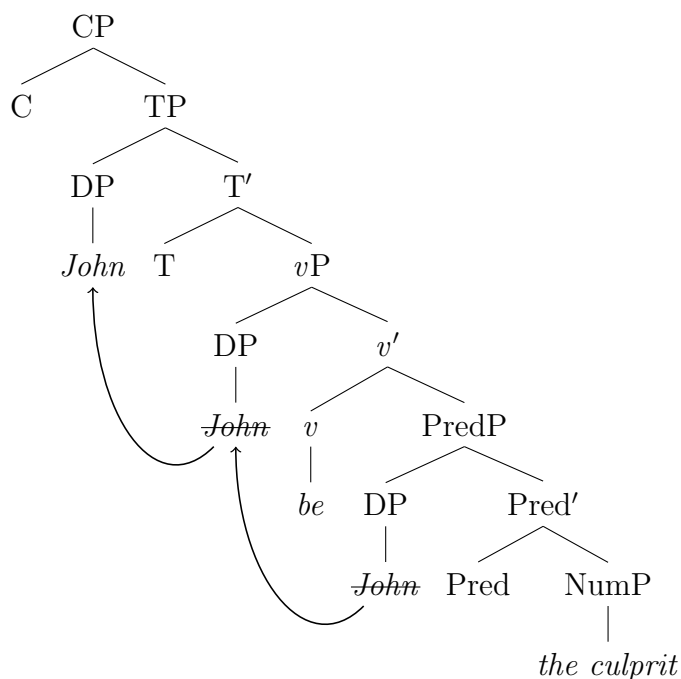
Table 8.1: ‘Specificational copular clauses’ Hartmann (forthcoming)

	<i>What’s new in the murder case? The culprit is John</i>	
	<i>John</i>	<i>the culprit</i>
Syntax	DP	DP
Focus	Focus	Background
‘Discourse’	New	Accessible

Table 8.2: ‘Predicative copular clauses’ Hartmann (forthcoming)

	<i>Do you think John or Bill is the culprit? John is the culprit</i>	
	<i>John</i>	<i>the culprit</i>
Syntax	DP	NumP
Focus	Focus	
‘Discourse’	Given	Given

(10)



8.3.2.1 Is Hartmann’s ‘predicative copular clause’ actually a reverse specificational?

Now, one important thing to immediately say about this is that it isn’t obvious to me that Hartmann’s ‘predicative copular clause’ is in fact a predicational copular sentence, at least not in Patten’s sense. In Patten’s sense, *John is the culprit*, understood predicationally,

should be telling us which class John is included in. But the context question Hartmann gives for her predicative copular clause *John is the culprit* (*Do you think John or Bill is the culprit?*) is rather asking for who is included in ‘the culprit’ class. As such, it looks superficially like a reverse specificational. For an example of a truly predicational *John is the culprit* (in Patten’s sense of predicational), the context would need to be something like:

- (11) New inspector on the case: There’s been a crime involving John, but I don’t know what his role in it was, whether he was the victim, the culprit etc.
 Officer who’s been on the case for a while: He/John’s the culprit.

I think here we can say—as Hartmann wants to about the cases she deals with—that both *John* and *the culprit* are ‘given’ by the context. But it would seem like the focus is on *the culprit* rather than *John*, and we would expect *the culprit* to receive nuclear tone, not *He/John*.

But Hartmann is explicit that her predicative copular clauses, whilst having focus on the subject, are not reverse specificationals. One argument she gives for this is from ‘plural tags’. The basic idea here is taken from Heycock (2012), for whom the plural tag data is intended to show that the specificational subject is not predicative because it is resumed in tags by a referential personal pronoun (*they*) rather than by *it*, which is taken to be the pronoun that is used to resume predicative elements in tags. This gives us the following contrast:

- (12) a. Blanchett and Nyong’o are the winners, although they don’t look it/*them.
 b. The winners were Blanchett and Nyong’o, *wasn’t it/weren’t they?

The evidence is problematic though. For a start they aren’t minimal pairs. If they were, with both having tag questions instead of just (12b), then we wouldn’t in fact see a contrast:

- (13) Blanchett and Nyong’o were the winners, *wasn’t it/weren’t they?

Equally, if we use the ‘look it’ tag of (12a), it seems to be fine with the other ordering of the flankers too:

- (14) The winners are Blanchett and Nyong’o, although they don’t look it/*them.

I think ‘look it’ is simply a fixed phrase that doesn’t allow for variation in the pronoun. So on this basis, I don’t think Hartmann has a strong argument here for saying that plural tags discriminate a ‘different semantic (and syntactic) status in the predicative copular clause order than in the specificational copular clause order’ (Hartmann, forthcoming, 19).

A second argument Hartmann gives for saying that her predicative copular clauses are not reverse specificationals is that ‘focus on the subject goes hand in hand with deaccenting

the post-copular material in the predicative copular clause ordering. Nothing of this sort is required in specificational copular clauses, where the respective noun phrase, *the culprit* [in *The culprit is John*], can or cannot be deaccented' (Hartmann, forthcoming, 19–20). This seems to be a fair description of the sentences in the contexts that Hartmann gives:

- (15) 'Predicative copular clause' question
- a. Do you think John or Bill is the culprit?
 - b. JOHN is the culprit./#JOHN is the CULPRIT
- (16) 'Specificational copular clause' question
- a. What's new in the murder case?
 - b. The culprit is JOHN./The CULPRIT is JOHN.

Only one flanker can have the nuclear tone of a sentence, so the capitalisations in (15b) and (16b) merely reflect some accenting as opposed to de-accenting. I'm not completely clear what Hartmann means here, though. There seem to be two possibilities: (1) that the difference between her two clause types is that focus on the subject in one type (the predicative copular clauses) necessitates de-accenting of the post-copula flanker in a way that it doesn't for the other type (the specificational copular clauses), or (2) that the difference is that the constituent that must be de-accented when there is subject focus in a predicative copular clause doesn't have to be de-accented in the specificational case. If she intended (1), then I think it's fairly clear that for the specificational case, (16b), there isn't any *option* with respect to the accenting of the post-copula flanker in (16); it can't be de-accented (i.e. **The CULPRIT is John*). So I think she intends (2).

But generally I think the pattern of accenting in these examples is an effect of the different givenness status of the flankers. In (15b) *John* must have contrastive focus, and *the culprit* is also given. As such, it would be odd to accent it. There would be no reason to (it is neither new information nor contrastive information). In fact elliding *the culprit* is perfectly possible here. In (16b), on the other hand, neither *the culprit* nor *John* are given. As such, both flankers could be accented to reflect the way they give new information (technically, Hartmann says *the culprit* is 'accessible' rather than new information). But the context question of (15a) seems like it could be perfectly felicitously answered with the specificational copular clause ordering:

- (17) a. Do you think John or Bill is the culprit?
- b. The culprit is JOHN./#The CULPRIT is JOHN./#The CULPRIT is John

As long as Hartmann agrees that (17b) really is a specificational sentence, then we see that

there is essentially the same requirement as in her ‘predicative copular clause’, (15), for de-accenting *the culprit*. In other words, her argument, re-stated as (2) just above (the constituent that must be de-accented when there is subject focus in a predicative copular clause doesn’t have to be de-accented in the specificational case) doesn’t hold. Just as in (15), *the culprit* must be de-accented in (17).

Hartmann states, then, that the sorts of examples we saw above when discussing Selvanathan’s approach are evidence for her approach and another argument for denying that her predicative copular clause and specificational copular clause are in fact both specificational. She gives the following examples, essentially of the same type as the examples in (7):

- (18) a. John, too, is Mary’s brother.
 b. Mr Clinton, too, is the candidate of ‘change’ (British National Corpus, AK9 1052)
- (19) a. #Mary’s brother is John, too.
 b. #Mary’s brother is also John.

She says: ‘These [‘focus-sensitive additive particles’ like *too* and *also*] cannot associate with the post-copula constituent in [specificational copular clauses]... Association is possible in [predicative copular clauses]’ (Hartmann, forthcoming, 20). But it isn’t obvious why this should be the case. It is a DP (*John, Mr Clinton*) that is focused in both cases (on Hartmann’s own analysis of her predicative copular clauses). The only potentially relevant difference I can see is that *John* ‘stays low’ in the specificational case for Hartmann i.e. it stays in Spec,PredP in (9), rather than rising to the Spec,TP canonical subject position. There may be other syntactic accounts that make the claim that for a constituent to associate with an additive particle like *too*, the constituent in question must occupy the canonical subject position of the clause, but Hartmann doesn’t make this claim herself, and it doesn’t seem immediately plausible, given the possibility of associating *too* with objects:

- (20) Mary kissed John and she kissed Bill too.

On the assumption that *too* can only associate with subjects in the canonical subject position, the sentences in (18) and (19) do become evidence for Hartmann’s syntactic analysis. Or at least, they become evidence for the part of it which has *John* in Spec,TP in (10) and staying low in Spec,PredP in (9). But that isn’t news; we knew already that *John* was the subject in *John is the culprit* but not in *The culprit is John* (where *The culprit* is the subject). And as I say, I doubt that the assumption (that *too* can only associate with constituents in

canonical subject position) would be warranted in any case.

In the next section, I attempt to account for the data in (7), (18), and (19) in terms of a necessarily exhaustive reading of the specificational subject that arises as a consequence of the inversion of the flankers in specificationals. But this is the only difference between them and reverse specificationals, which simply don't require the attributive flanker to have an exhaustive reading.

So in this section, I've looked at the arguments that Hartmann gives to claim that her predicative copular clauses are not reverse specificationals. None of these arguments seem problem-free, and so for now I'll hold on to the view that they are reverse specificationals. But I think Hartmann's work certainly demonstrates the need for finer distinctions to be made. There seem to be at least three different kinds of context question with different patterns of acceptable accenting:

- (21) Specificational context question
- a. Is John or Bill the culprit?
 - b. The culprit is JOHN. / #The CULPRIT is John. / #The CULPRIT is JOHN.
JOHN is the culprit. / #John is the CULPRIT. / #JOHN is the CULPRIT.
- (22) Predicational context question
- a. Is John the culprit or the victim?
 - b. #The culprit is JOHN. / #The CULPRIT is John. / #The CULPRIT is JOHN.
#JOHN is the culprit. / John is the CULPRIT. / #JOHN is the CULPRIT.
- (23) Third type of context question
- a. What's new in the murder case?
 - b. (?)The culprit is JOHN. / #The CULPRIT is John. / The CULPRIT is JOHN.
#JOHN is the culprit. / #John is the CULPRIT. / JOHN is the CULPRIT.

As indicated above, I think part of the explanation of this pattern of acceptable accenting lies in the givenness status of the flankers. Below, though, I will simplify matters and ignore the third type in (23) above.

8.3.2.2 Conclusions about Hartmann's analysis

So whilst the approach in Hartmann (forthcoming) is very much in the spirit of the kind of description which this thesis has tried to argue will be successful, there are a few details to note about it. Firstly, as we've just seen, it isn't completely obvious that her predicative copular clauses are what others call predicational copular sentences and that her specificational

copular clauses are what others call specificational copular sentences (more so than a ‘third type’, as in (23) above). This isn’t a problem, as such; it’s just important to be clear about the differences. Secondly, we haven’t really motivated the way that *the culprit* is sometimes a DP, and sometimes a NumP. Thirdly, it isn’t clear to me what exactly the nature is of the ‘type mismatch’ which causes a ‘Focus–Background’ mapping, which then assigns a formal feature that drives inversion. Fourthly, I’m not sure what the focus status of *the culprit* is in Hartmann’s ‘predicative copular clause’. She doesn’t explicitly say that it is the background (hence the gap in Table 8.2), but if it isn’t the background, then why not? What else could it be? From what we said about focus in 7.6.3.1, we saw that the ‘background’ of a sentence is a set of alternative linguistic expressions, and so not strictly speaking a component of any given sentence. For the context questions that Hartmann gives

- (24) a. What’s new in the murder case? (specificational question)
 b. Do you think John or Bill is the culprit? (predicational/reverse specificational question)

I think the alternatives to (24a) would be expressions like *The victim’s family have been informed/Some fingerprints were found on a door handle/The culprit is John etc.* They are things that could be ‘new in the murder case’. So it isn’t clear that it is right to designate *the culprit* as ‘the background’, as Hartmann does. In fact it isn’t obvious that the focus of Hartmann’s specificational wouldn’t be the whole response to (24a) (i.e. *The culprit is John*, rather than just *John*). This, incidentally, may be a way in which we could distinguish the ‘third type’ that we identified at the end of the last section, example (23). The background of (24b) is the set of all x such that x is the culprit, where this has been restricted to exactly two members, John or Bill. Selecting John from this set makes it the focus of *John is the culprit*, and it is more obvious that *the culprit* could be called the background in this case, as every alternative sentence contains this constituent.

As I say, Hartmann’s (forthcoming) analysis is very much in the spirit I want to pursue here in terms of understanding the difference between pairs like (1) and (2). But I think these details just noted mean that we should develop it slightly differently.

8.4 An alternative analysis

8.4.1 Exhaustivity

So I shall assume that *John* is focused in both (18a) and (19a), and *too* should be able to associate with a focused constituent. But as Hartmann (p.c.) notes ‘the focus properties

seem to be different’. This is the tack which I shall now pursue.

To re-cap, the problem for me is this: I want to see if I can keep the syntax and semantics of these sentences ((1) and (2)) identical (modulo the word order difference; see below, section 8.4.3, for a reason *why* this word order difference may be necessary).

We’ve seen a distinction that Patten (2012; 2016) makes between ‘predicational’ and ‘specificational’ readings in terms of which flanker is focused. Where the QUD is (25a)—the predicational reading—then we say that focus is on *the culprit*. When the QUD is as in (25b)—the specificational reading—then we say that focus is on *John*. In set-theoretic terms, this could be put as:

- (25) a. ‘For a given $x_1 \dots x_n$, what is the P, such that $x_1 \dots x_n \in P$?’ = predicational
e.g. ‘What set is John a member of?’
b. ‘For a given P, what are the $x_1 \dots x_n$ such that $x_1 \dots x_n \in P$?’ = specificational
e.g. ‘Who or what are the members of the set of culprits?’

That’s the first difference, then. The second distinction we made was between the way that the flankers refer. We said that *the culprit* is used attributively (in the sense of Donnellan (1966)), where *John* is used referentially. We can extend this to the problematic examples with possessives and say in examples like (7), *John’s friend* refers attributively to *whoever* is the person who is John’s friend.

So in the problematic examples like (7), we are dealing with cases where the focus is on *Susan/Mary* in both orderings (7a) and (7b). (7a), then, is here what the literature refers to as a ‘reverse specificational’. How do we account for the way that (7a) is ok with *too* and (7b) isn’t? Well, as we saw, Selvanathan (2016) offers us a semantics that disambiguates them. But I don’t think he really solves the problem. In terms of set membership relations, his semantic representations in (8) are identical, giving us the following sentence semantics ((26a) corresponds to (8a) and (26b) corresponds to (8b)):

- (26) *Susan is John’s friend*
a. λx [friend-of (x, John)] (Susan) \rightarrow friend-of (Susan, John)
b. λy [ιx . friend-of (x, John) & x = y] (Susan) \rightarrow ιx . friend-of (x, John) & x = Susan

All that differentiates them, then, is the uniqueness requirement of (26b), which states that there is one and only one friend of John. Now, this may well go to the core of what the problem is with the examples in (7), (18) and (19), but it isn’t clear to me that this is doing any more work than re-stating the problem. If we say that the problem with (19a) is that *Mary’s brother* can only stand for a maximum of one brother, then we can see how an

additive particle, or even the second conjunct without the additive particle, is not going to make sense. The question, then, is why this is imposed in (19) but not (18)? Selvanathan's solution is to introduce the iota operator. The ultimate reason for this, as noted above, seems to be that his specificationals (i.e. (1) orderings) *must* be equations. But note that he could have an equation with an existential operator, so his rule whereby specificationals must be equations doesn't necessarily require the iota operator to scope over the object denoted by *John's friend* i.e.

(27) $\exists x$. friend-of (x , John) & $x = \text{Susan}$

My preferred approach to this problem is to relate it to a property of the information structure of these sentences. Let's take the original problematic examples:

(28) Susan is John's friend, and Mary is John's friend too. = (7a)

(29) */?John's friend is Susan and John's friend is Mary too. = (7b)

We can say that in (29), the first conjunct has 'exhaustive focus' on *Susan* (for exhaustivity as an information structural notion, see (Horn, 2016, 117ff.)). What this means is that the set of alternatives corresponding to the QUD in (30) with every possible value for x is exhausted by the value 'Susan':

(30) Who is the x such that x is John's friend?

By stipulation then (as it has been throughout), when the order is *John's friend is Susan*, *Susan* has exhaustive focus. The difference with *Susan is John's friend* is that *Susan* there does not have to have exhaustive focus. Susan can merely be one of the possible values for x in (30); there may be others (Mary, Bill etc.). Taking Hartmann's version of the problematic sentences in (18) and (19), we see how this difference shows up if we adjust the discourse contexts for the sentences, such that the presupposition of exhaustive focus on *John* when *Mary's brother* precedes *John* clashes with the presupposition of the discourse context:

(31) A: So let's talk about Mary's relatives. We know that Mary has several brothers, but who are they?

B: Well, JOHN is Mary's brother, for a start, and STEVE TOO is Mary's brother.

B': Well, */?Mary's brother is JOHN, for a start, and Mary's brother is STEVE TOO.

(32) A: So let's talk about Mary's relatives. We know she has one brother and two sisters, but who are they?

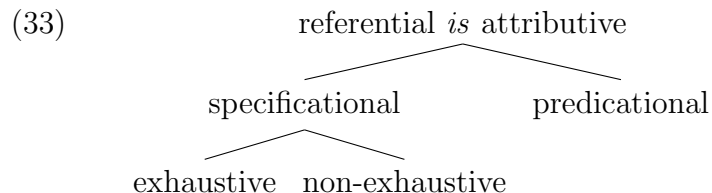
B: Well, JOHN is Mary's brother, and SUSAN and GLADYS are her sisters.

B': Well, Mary's brother is JOHN, and her sisters are SUSAN and GLADYS.

So in a discourse context where Mary has multiple brothers, we can produce the *John/Steve is Mary's brother* order, along with an additive *too* after the focused constituent in the second conjunct. But this is not possible for the *Mary's brother is John/Steve* order, (31B'); it doesn't make sense to exhaust the value of the x that is Mary's brother with one value when the discourse context tells you that one value cannot possibly be exhaustive. And then of course it doesn't make sense either to introduce a second conjunct, a second value for a variable that has been exhausted by *John* in the first conjunct. In (32) the discourse context is that Mary has one and only one brother (and two sisters). (32B') exhausts this value, which is not inconsistent with the discourse context. There is no clash between the discourse context and the exhaustive focus on *John* in (32B'). As I said, the *John* of *John is Mary's brother* doesn't necessarily have exhaustive focus. As such *John is Mary's brother* in both (31B) and (32B) is compatible with either discourse context. If there are multiple brothers, as in (31), then *Mary's brother* is simply one of them. If there is only one brother in the context, then *Mary's brother* is that one brother.

8.4.2 The overall picture

The overall picture that we arrive at, then, is as in Table 8.3 (where I state things in terms of set membership again, but note that this is merely for exposition), or alternatively as follows⁹:



⁹I haven't explicitly pointed out that 'referential *is* attributive' allows for an exhaustive as well as a non-exhaustive specificational reading. But I assume it is from the felicity of examples like B's response in (32). I should also admit here that I haven't even looked at the properties of exhaustivity relative to the attributive element when it is focused in predicational structures. There is certainly something to say about this which would expand the description I've given here.

Table 8.3: Three kinds of copular sentence

<u>Specificational—exhaustive</u> QUD: For a given P, what is the one and only x such that $x \in P$? / what are all $x_1 \dots x_n$ such that $x_1 \dots x_n \in P$?	<i>The culprit is John, John is the culprit</i>
<u>Specificational—non-exhaustive</u> QUD: For a given P, what are the $x_1 \dots x_n$ such that $x_1 \dots x_n \in P$?	<i>John is the culprit</i>
<u>Predicational</u> QUD: For a given $x_1 \dots x_n$, what is the P, such that $x_1 \dots x_n \in P$?	<i>John is the culprit</i>

- (34) attributive *is* referential
 |
 specificational
 |
 exhaustive

This, incidentally, is not the same as the way the literature normally captures the specificational vs. predicational distinction. Arche et al. (forthcoming, 37) say ‘most theories’ capture the difference in terms of:

- (35) a. Predicational: [More referential NP] is [Less referential NP]
 b. Specificational: [Less referential NP] is [More referential NP]

I think the core of the difference between specification and predication is rather to do with which flanker is focused. Also, the notion ‘more/less referential’ is not well-defined. Here it is specifically the notion of attributive or referential use of an expression, something which isn’t the same as a gradable notion of referential specificity (if that’s what’s intended). Furthermore, it fails to recognise the empirical fact of reverse specificationals, where the ‘more referential NP’ actually precedes the less referential one.

So I have not thus far encountered a need to distinguish (1) and (2) semantically or syntactically (modulo the word order difference, see section 8.4.3 for why this arises). For Patten (2012; 2016), there is a ‘semantic relation’ of ‘class inclusion’ that is the same between specificational and predicational copular sentences. Personally, I don’t think we need to stipulate a semantic relation as long as we recognise that there is a crucial difference between

the flankers in terms of how they refer: one of them is attributive and the other is referential.¹⁰ Patten makes ‘inclusiveness’ a property of definites (see the end of section 8.2).¹¹ This seems to be right. So we don’t find the same ambiguity with attributively used definite descriptions as we do with possessives:

(36) *The culprit is BRIAN and the culprit is MARY TOO.

The difference between a specificational sentence and a predicational one, then, seems to follow here from focus considerations. Where the focus is on the referential flanker, then the sentence is specificational. Where the focus is on the attributive flanker, then the sentence is predicational.

8.4.3 Alleviating stipulations

There are stipulations in this account which can be alleviated. I won’t go into the details here, but a consensus is emerging that:

(37) referential_{focus} *is* attributive

is the default word order, such that

(38) attributive *is* referential_{focus}

is non-default, or an ‘inversion’ of the attributive flanker (in a derivational approach) (Heycock, 2012). As I said in section 7.5.5.2, inversion per se can be seen as enabled by a property of UG (the possibility of syntactic displacements due to Internal Merge). However, the reason *why* a given constituent will move is the province of information structure. The information structure properties are well studied (see e.g. Birner (1994), Birner et al. (2007)). One ‘robust pragmatic constraint’ Patten (2016) draws attention to is Birner (1996, 90)’s: ‘[T]he preposed element. . . must not be newer in the discourse than the postposed element’. As we saw in section 7.6.3.2, the given–new distinction is different to the focus–background distinction. However, if we take the classic specificational context question *Who’s the culprit?* (at the scene of a crime) then we can see that the focus of the reply to that question (the item selected from the set of alternatives, every x such that x is the culprit) will necessarily be newer information than *the culprit*, which Hartmann (forthcoming) calls the ‘background’.

¹⁰A potential problem in not stating a relation between the flankers is the ‘unity of the proposition’, the problem of how it comes about that a sentence is a unity with a truth value and not a mere list of words. Of course, I do have a way of putting the flankers of a copular sentence together; I said in section 7.5 that copular sentences are all consistently *syntactic* predications. As such, what I say is at least not inconsistent with Collins’s (2011) and Reichard’s (2013) grammatical solutions to the unity problem.

¹¹It isn’t clear to me whether inclusiveness is the same phenomenon as exhaustivity or not.

As such, *the culprit* can be preposed. If things were the other way, and the *the culprit* was the focus, it would be newer information, and it would not be possible to prepose it. So the fact that *the culprit* is preposed in (1)—*The culprit is John*—logically implies that it cannot be the focus, and therefore that *John* must be.

To put this more simply, we can understand that inverting *the culprit*, the attributively used flanker which by default is post-copula, will engender that the referentially used flanker is focused. If we strengthen this to say that this inversion has the effect of putting exhaustive focus on the referential flanker, then from this three things that have been stipulated so far follow:

1. ‘attributive *is* referential’ can only be exhaustive specificational, and so logically
2. ‘attributive *is* referential’ can only be specificational
3. If one isn’t dealing with a focus that can be ambiguous between an exhaustive or non-exhaustive reading (where the attributive flanker is a definite description, for example) then there should be free variation in the ordering of the flankers i.e. the reason why we can say *The culprit is JOHN* or *JOHN is the culprit* in response to *Who is the culprit?* is because we don’t need to order the flankers (as we do with possessives like *Mary’s brother*) in order to signal if we are using them exhaustively or not.

8.4.4 Extension to pseudoclefts

Den Dikken gives the following example:

- (39) What Brian doesn’t eat is food for the dog. (example taken from den Dikken (2006a, 269))

Den Dikken’s explanation for the ambiguity of this example is to say that there are two readings corresponding to (what he understands as) the specificational and predicational distinction:

- (40) a. Brian doesn’t eat the following: food for the dog—specificational
 b. The stuff that is left over on Brian’s plate is fed to the dog, serves as dog food—predicational

By now, it should be clear that tagging readings in this way doesn’t get us very far in describing their differences, so let’s briefly see how the approach here would deal with it.

Firstly, there just seems to be a straight lexical ambiguity: *what Brian doesn’t eat* can either mean (1) the things that Brian doesn’t like eating, his eating dislikes, (2) the things

that are left on Brian's plate i.e. his leftovers. Equally *food for the dog* is also ambiguous between (1) what people would more idiomatically call *dog food* i.e. commercial dog food that you can buy in a shop, (2) the stuff that a particular dog gets given to eat by its owner. It's odd that den Dikken seems to try to explain a lexical ambiguity on the basis of what is for him a syntactic distinction between a sentence's being predicational and its being specificational. It's not clear that lexical ambiguities can ever be explained by syntactic facts (unlike a real syntactic ambiguity like *Brian saw the man with a telescope*). We can notate the lexical ambiguity:

- (41) a. [[Brian's eating dislikes]] are [[commercial dog food]]
 b. [[Brian's leftovers]] are [[what the dog gets given to eat]]

Now, the next parameter we've identified in this chapter is the way that these flankers refer. I think it's safe to say that they both allow for an attributive and a referential use. *Brian's eating dislikes* attributively are just *whatever* Brian doesn't like eating; referentially, they can be given extensionally e.g. {cabbage, beetroot, gherkins}. And so on for the others. So if we just take (41a), we have the following possible combinations:

- (42) a. [[Brian's eating dislikes]]^{referential} are [[commercial dog food]]^{referential}
 b. [[Brian's eating dislikes]]^{referential} are [[commercial dog food]]^{attributive}
 c. [[Brian's eating dislikes]]^{attributive} are [[commercial dog food]]^{referential}
 d. [[Brian's eating dislikes]]^{attributive} are [[commercial dog food]]^{attributive}

Finally (I'll ignore exhaustivity here for the sake of exposition), either flanker can be the focus, which I'll mark in bold. If we just take (42c) as an example:

- (43) a. [[Brian's eating dislikes]]^{attributive} are [[**commercial dog food**]]^{referential}
 b. [[**Brian's eating dislikes**]]^{attributive} are [[commercial dog food]]^{referential}

Now, as we've seen, the attributive cannot be focused when it precedes the referential, so (43b) should really be out, which does seem to be the case:

- (44) A: I'm not sure if food for the dog is what Brian doesn't eat or what Bill doesn't eat.
 B: ?**What Brian doesn't eat** is food for the dog.

Overall then, on the basis of the two properties of attributive/referential and focus, I predict eight possible readings for each extension pair, one of which—where the attributive flanker precedes the referential one and is intended to satisfy the QUD, as in (44b)—is infelicitous. Showing that that all these readings are available and follow the same QUD felicity constraint

Table 8.4: *Brian, the culprit*: possible permutations

1	the culprit ^{referential}	Brian
2	the culprit ^{referential}	Brian
3	the culprit ^{attributive}	Brian
4*	the culprit ^{attributive}	Brian
5	Brian	the culprit ^{referential}
6	Brian	the culprit ^{referential}
7	Brian	the culprit ^{attributive}
8	Brian	the culprit ^{attributive}

as the simple cases is evidence for the approach here. This is what I'll now do, but I'll revert to the simple case for the exposition, before coming back to these pseudocleft cases. The point of this section has just been to show that the approach here can extend straightforwardly to pseudoclefts, and in particular that this is explanatory and (will be shown to be) more empirically adequate than den Dikken's characterisation of the ambiguity of *What Brian doesn't like is food for the dog* in terms of two readings corresponding to the specificational and predicational ones.

8.4.5 A more pluralistic taxonomy

So let's take the simple case:

- (45) a. The culprit is John.
b. John is the culprit.

There are two crucial parameters: (1) the way of referring: referential vs. attributive (2) which flanker is focused. In the simple case, *John*—as a proper name—is not ambiguous between a referential or attributive way of referring; *the culprit* is though. This then gives us eight logically possible permutations of the two orderings in (45), see Table 8.4. As I did in the last section 8.4.4, I notate the way of referring with superscripts and focus with bold.

Now, one of these permutations is bad (marked with an asterisk): line 4, where a focused attributive precedes a non-focused referential (which is generally a stipulation, though see 8.4.3 for an attempt to explain why this is the case). Are these seven sentences actually there? I believe so. Four of them (lines 1, 2, 5, and 6) are equatives, and so would be uttered in contexts like:

- (46) a. I saw 'the culprit' yesterday. And this morning I saw Brian. They look a lot

alike.

- b. No you're mistaken, Brian IS the culprit. They aren't two different people as you seem to think.

Does focus apply to the flankers of equatives? I mentioned earlier (footnote 7) that Hartmann and Hegedűs (2009) believe that Hungarian equatives can exhibit focus–background structure, though it isn't clear to me exactly how this would be the case. Perhaps a context like the following would be more clearly one where one of the flankers is focused:

- (47)
- a. I know my male friend who I call 'the culprit' is identical to someone else, but I'm not sure who it is.
 - b. He's BRIAN.

So let's assume focus can apply to flankers of equatives. We might just need a different analysis for the cases like (46), which after all does appear to be the more standard context of a typical identity sentence.

The other three readings are: line 3 correlates with a normal specificational sentence, line 7 with a normal predicational sentence, and line 8 with a 'reverse specificational'.

Let's move on to the pseudocleft case, then, which is only more complicated than the simple case because here *both* flankers can in theory be either referential or attributive. We therefore predict that there will be double the number of readings as we saw for the simple case. Instead of 7 readings, we predict there to be 14. Given our example *What Brian doesn't eat is food for the dog* was lexically ambiguous, there should in fact be 14 readings per lexical pair combination, so 28 readings given that there were two lexical pair combinations ('Brian's eating dislikes + commercial dog food' and 'Brian's leftovers + what the dog gets given to eat'). Let's just pretend that it is not a lexically ambiguous pseudocleft and just articulate the 14 readings for the 'Brian's eating dislikes + commercial dog food' lexical pair. Again, this can be presented in a table, as in Table 8.5.

There are two unacceptable combinations this time, lines 6 and 14, representing the impossible focused attributive preceding a non-focused referential. This time, there are eight equatives, line 1, 2, 7, 8, 9, 10, 15, and 16. I noted some reservations about these readings above in relation to whether these can be focus–background structures as I assume. Here, lines 7, 8 and 15, 16 are combinations of two attributive flankers. Is it right to call these equatives? I think it is. Essentially, they would mean:

- (48) Whatever Brian's eating dislikes are are the same things as whatever commercial dog food is.

Table 8.5: *What Brian doesn't eat* (= 'Brian's eating dislikes', ED), *food for the dog* (= 'commercial dog food', DF): possible permutations

1	ED ^{referential}	DF ^{referential}
2	ED ^{referential}	DF ^{referential}
3	ED ^{referential}	DF ^{attributive}
4	ED ^{referential}	DF ^{attributive}
5	ED ^{attributive}	DF ^{referential}
6*	ED ^{attributive}	DF ^{referential}
7	ED ^{attributive}	DF ^{attributive}
8	ED ^{attributive}	DF ^{attributive}
9	DF ^{referential}	ED ^{referential}
10	DF ^{referential}	ED ^{referential}
11	DF ^{referential}	ED ^{attributive}
12	DF ^{referential}	ED ^{attributive}
13	DF ^{attributive}	ED ^{referential}
14*	DF ^{attributive}	ED ^{referential}
15	DF ^{attributive}	ED ^{attributive}
16	DF ^{attributive}	ED ^{attributive}

It's just a kind of equative of unknown terms; the speaker doesn't know either what Brian's eating dislikes are or what commercial dog food is, he just (somehow) knows that whatever they are they are the same thing. The other six readings are: lines 5 and 13 correlate with normal specificational sentences; lines 3 and 11 correlate with predicational sentences, and lines 4 and 12 correlate with reverse specificationals.

We can apply context questions in order to check that these readings are available. For example, for the normal specificationals of lines 5 and 13:

(49) Line 5—normal specificational

A: There are certain things that Brian doesn't eat, can you give me an example?

B: What Brian doesn't eat is food for the dog (i.e. Brian's eating dislikes include commercial dog food like Pedigree Chum).

(50) Line 13—normal specificational

A: There are certain things that are food for the dog (i.e. commercial dog food). Can you give me an example?

B: Food for the dog is what Brian doesn't eat (i.e. The stuff that you know Brian doesn't eat, Pedigree Chum, is an example of what commercial dog food is)

For the predicational sentences of lines 3 and 11:

(51) Line 3—predicational

A: I've heard that Brian's eating dislike is something called 'Pedigree Chum'. But I don't know what it is, so I don't know what what Brian doesn't eat is. I think it's either dog food or cat food.

B: What Brian doesn't eat is food for the dog (i.e. Brian's eating dislike/Pedigree Chum is commercial dog food).

(52) Line 11—predicational

A: I know what commercial dog food is; it's stuff like Pedigree Chum. But can you tell me something else about it?

B: Food for the dog is what Brian doesn't eat. (i.e. Stuff like pedigree chum is (amongst) Brian's eating dislike(s))

For the reverse specificationals of lines 4 and 12:

(53) Line 4—reverse specificational (equivalent to line 13)

A: There are certain things that are food for the dog (i.e. commercial dog food). Can you give me an example?

B: What Brian doesn't eat is food for the dog (i.e. The stuff that you know Brian doesn't eat, Pedigree Chum, is an example of what commercial dog food is)

(54) Line 12—reverse specificational (equivalent to line 5)

A: There are certain things that Brian doesn't eat, can you give me an example?

B: Food for the dog is what Brian doesn't eat (i.e. Brian's eating dislikes include commercial dog food like Pedigree Chum).

The two illicit readings then would correspond to the contexts (capitalization marks focal intonation):

(55) Line 6—illicit predicational with focused attributive preceding referential

A: I know what commercial dog food is; it's stuff like Pedigree Chum. But can you tell me something else about it?

B: #WHAT BRIAN DOESN'T EAT is food for the dog. (i.e. Stuff like pedigree chum is (amongst) Brian's eating dislike(s))

(56) Line 14—illicit predicational with focused attributive preceding referential

A: I've heard that Brian's eating dislike is something called 'Pedigree Chum'. But I don't know what it is, so I don't know what what Brian doesn't eat is. I think it's

Table 8.6: *The winner of the election might have been the LOser*—eight readings (Higgins 1979, 271–3)

Identity	1. referential + referential
	2. attributive + referential
	3. referential + attributive
	4. attributive + attributive
Specificational	5. superscriptional + attributive specificational
	6. superscriptional + referential specificational
Predicational	7. attributive + predicational
	8. referential + predicational

either dog food or cat food.

B: #FOOD FOR THE DOG is what Brian doesn't eat (i.e. Brian's eating dislike/Pedigree Chum is commercial dog food).

These judgements certainly aren't as clear as in the simple case:

(57) A: Is John the culprit or the victim of the crime?

B: #The CULPRIT is John.

The reason, I think, is that we have a natural tendency to 'hear' as acceptable a string of words which strictly speaking are unacceptable in the context, if that string of words can be acceptable in another possible context (though this doesn't explain why (57) is so relatively clearly bad). The result is confusion, but not a very clear intuition about acceptability, unfortunately; it's possible that there are ways to render the acceptability judgements clearer.

The point, once again, is to say that these permutations are all essentially distinct sentences. There are 14 distinct sentences here, 7 for each ordering of the flankers.

8.4.6 Returning to Higgins's eight-ways ambiguous example

At the end of the introduction to this thesis, section 1.3, I gave an example that Higgins gives himself of an eight-ways ambiguous example with the readings as in Table 1.5, repeated in 8.6.

(58) The winner of the election might have been the LOser.

Let's analyse this now within the system just proposed. The first thing to realise is that this sentence is like the pseudocleft case in allowing *both* flankers to refer referentially or

attributively. With one ordering, then, we predict that there should be eight readings, one of which will be illicit.

Higgins claimed that there were four identities, lines 1 to 4 of Table 8.6. My system also predicts there are four identities. However, these are only where both flankers are referential or both attributive. So what can I say about Higgins's lines 2 and 3? Bearing in mind his fixing the reference of the terms as Nixon for the referential use of *the winner of the election* and McGovern for the referential use of *the loser of the election*, we should have the following possible readings:

- (59) a. Line 2: [The winner of the election]^{attributive} might have been [the LOser]^{referential}
 = Whoever won the election might have been the same person as McGovern.
 b. Line 3: [The winner of the election]^{referential} might have been [the LOser]^{attributive}
 = Nixon might have been the same person as whoever lost the election.

I'll reserve judgement on what to say about the examples in (59) for the moment. If they are there, then they would add two readings to the $8(-1 = 7)$ that I would predict for this sentence.

Higgins's lines 5 and 6 are two specificationals with the difference that the predicate is attributive or referential, respectively. I don't have a 'superscriptional' category to define the subject of a specificational sentence myself; for me, this category is essentially a non-focused (or background) attributive flanker. So Higgins's line 5 would be (bold marking focus) an 'attributive + **attributive**' combination, and his line 6 would be 'attributive + **referential**'. Line 6, I agree, is a normal specificational. But for the 'attributive + **attributive**' of line 5, I have so far been thinking of this as an equative, moreso than a specificational. The reading Higgins gives for his line 5 can be paraphrased:

- (60) The one who might have won the election was whoever was the loser.
 = The following person might have won the election: whoever lost it.

This certainly seems like a possible reading too. As such, I may have to consider a further extension to my system. This is a serious challenge as I have hitherto been considering focus on the attributive flanker (at least when the other flanker is referential) as corresponding to a predicational reading, not a specificational one, as indeed appears to be possible here.

The other two readings in Higgins's system, lines 7 and 8, are two predicational readings. I don't have a 'predicational' category of flanker; this rather corresponds to a focused attributive flanker for me. As such, Higgins's line 7 would apparently correspond to an 'attributive + **attributive**' combination for me. There are problems with this. This is exactly the same combination that I just correlated with Higgins's line 5 specificational (which I had

Table 8.7: *The winner of the election might have been the LOser*: possible permutations

Permutations			Higgins's
1	the winner ^{referential}	the loser ^{referential}	1(?)
2	the winner ^{referential}	the loser ^{referential}	1(?)
3	the winner ^{referential}	the loser ^{attributive}	8 (predicational)
4	the winner ^{referential}	the loser ^{attributive}	? (reverse specificational)
5	the winner ^{attributive}	the loser ^{referential}	6 (specificational)
6*	the winner ^{attributive}	the loser ^{referential}	?
7	the winner ^{attributive}	the loser ^{attributive}	4(?)
8	the winner ^{attributive}	the loser ^{attributive}	4(?)

thus far been thinking of as an equative in any case). Higgins paraphrases his line 7:

(61) Whoever won the election might have lost it.

Is this really distinct from (60)? I have no reason to doubt it. This, again, is a serious challenge to me, as it seems to indicate that there is something to a predicational reading that goes beyond merely focus on an attributive flanker, as I've been assuming so far (in the context of examples where the other flanker is referentially used). Higgins's line 8, on the other hand, corresponds straightforwardly to a 'referential + **attributive**' combination, for me, which I've also said can be seen as a predicational sentence. Table 8.7 presents the analysis that I would give for Higgins's ambiguous sentence and the correspondence with Higgins's categories in the final column. As we can see, it is seriously incomplete. Lines 2, 3, 5, and 7 of Higgins's Table 8.6 find no correspondent in my system. Lines 4 and 6 of my table seem to have no correspondent in Higgins's table (although line 6 is an unacceptable focused attributive preceding a background referential flanker, so Higgins may simply have decided not to mention it).

8.4.7 Conclusions in the light of the comparison with Higgins

I'll just say a few words here on how the analysis I propose should be adapted in the light of the comparison with Higgins's eight-ways ambiguous example. An important issue that I've overlooked is the nature of equatives. An equative does not simply arise where both flankers are referential or both attributive. Higgins's examples seem to clearly show that there can be equatives where one of the flankers is referential and the other attributive (his lines 2 and 3 of Table 8.6). I accept this without much argument. I expressed suspicions in 8.4.5 about whether we can think of there being four different kinds of equative, as in Table 8.4,

partly because it isn't clear how focus applies in equatives. One obvious thing to say about equatives is that nuclear tone is often on the copula itself:

- (62) Exterior of Monk's coffee shop. Cut to Jerry and George at their regular booth.
 George: What is Holland?
 Jerry: What do you mean, 'what is it?' It's a country right next to Belgium.
 George: No, that's the Netherlands.
 Jerry: Holland IS the Netherlands.
 George: Then who are the Dutch?
 (from TV series *Seinfeld* series 9, episode 1 'The Butter Shave')

As such, we might consider that the background of an equative are two alternatives, paraphrasable as: *Holland and the Netherlands are distinct* and *Holland and the Netherlands are not distinct*. Jerry's utterance of *Holland IS the Netherlands* selects the latter alternative. Thus, it is the whole sentence that Jerry utters, rather than any individual flanker, that is the (contrastive) focus of an equative. Given this, we can see how equatives differ from specificational and predicational, and also how they can in theory manifest all permutations of the referential/attributive distinction.

Furthermore, it seems that there can be specificational and predicational sentences where the flankers are identical with respect to the referential/attributive distinction (Higgins's lines 5 and 7). This does require me to change a core tenet of my system so far: namely, that 'specificational' is defined as 'focus on the referential flanker' and 'predicational' is defined as 'focus on the attributive flanker'. As noted above already, I suspect this is only a fair description when the other flanker is attributive or referential, respectively. Now, I can't stop there and just say that 'attributive + **attributive**' is not, as I previously thought, an equative, but rather predicational. This is because both lines 5 and 7 are 'attributive + **attributive**' and 5 is manifestly specificational. So there really must be something else that goes beyond the attributive/referential distinction and focus but which marks an apparently specificational vs. predicational distinction. I can retain the generalisation about the difference between specificational and predicational readings when the flankers have a different referential/attributive status, but I do have to say something special about the case where we have two attributive flankers. It isn't obvious to me what the right approach would be here i.e. how to distinguish:

- (63) *The winner of the election might have been the LOser.*
 a. The one who might have won the election was whoever was the loser = The following person might have won the election: whoever lost it.—specificational

- b. Whoever won the election might have lost it.—predicational

There seems to be a difference with respect to the scope of the modal verb which may be crucial. In the specificational case, it relates to the *possible winner* (if things had been different)—*the winner* in the (63a) reading did not in fact win the election. In (63b) however, the modal relates to a *possible loser*; *the winner* there did in fact win the election. Whether this difference (in the scope of the modal verb) suffices to distinguish (63a,b) or whether there may need to be other adjustments, I don't know. Of course, if there is a need for further adjustments, then my instinct would be to look at the information structural properties of the two sentences in more depth. Related to the difference in the modal verb, there will be a difference in the alternatives that make up the background of (63a,b). In (63a) the only alternative is the 'actual loser'. In (63b) the background alternative is the 'possible loser' of the election. So they at least have different focus properties on this basis.

8.5 Arregi et al. (2018)—Arguments that (1) and (2) are not equivalent in their formal semantics

Arregi et al. (2018) have recently argued explicitly in defence of the view that (1) and (2) should not be considered truth-conditionally equivalent. And so the fundamental difference between them relates to a difference at the level of their formal semantics. This position is directly at odds with what I say here, so I shall briefly run through the two arguments they give (Arregi *et al.*, 2018, section 5) and my replies to them.

The first argument relates to a semantically substantive difference in the two flanker orderings when the the sentence is negated:

- (64) a. One of the prices is not 3.99.—Specificational: one-of-the-prices $>$ \neg
 b. 3.99 is not one of the prices.—Predicational: $\neg >$ one-of-the-prices

With respect to the list of prices in (65), they point out that (64a) would be true where (64b) would be false:¹²

¹²Given that negation is generally considered to take scope at the level of Logical Form (as well as S(urface)-Structure) in theories of Universal Grammar, it should perhaps be expected that (64a,b) would be ambiguous between a reading where 'one of the prices' scopes over negation and a reading where negation scopes over 'one of the prices' i.e. equivalent to 'It is not the case that one of the prices is 3.99'. This would be expected as long as we don't consider these sentences to contain any structure that would block LF-movement (which is always invisible). This fact would also undermine the reply to this evidence that I give below. However, it isn't obvious to me that (64a,b) is ambiguous; the S-Structure position of negation does indeed seem to reflect the only possible c-command (and hence scope) relation in these sentences.

- (65) *List of prices at a store*
 flour: \$3.99 sugar: \$3.99
 bread: \$3.99 milk: \$2.00
 salt: \$3.99 pepper: \$3.99

(65a) is true because it says that there is one price (which there is, the price of milk) which is not \$3.99 (it's \$2.00), and (65b) is false because it says that \$3.99 is not amongst the prices on the price list, but in fact it is there five times. This comes out semantically in Arregi et al.'s system as:

- (66) a. There exists a price concept in the world of evaluation which does not equal 3.99.—Individual concept/specificational
 b. No price is 3.99.—predicational

According to Arregi et al. (66b) is the only meaning that the inverse predication theory makes available. This is indeed the prediction of a theory that starts with an analysis of (64a) as an $\langle e, t \rangle$ function, comparable to '3.99 is not a price'. The problem with Arregi et al.'s analysis is that it doesn't say *why* we should assume one analysis for the semantics of *one of the prices* when it is post-copula and another analysis when it is pre-copula. And so we don't actually know why there shouldn't be a reading for (64b) along the lines of (66a), which of course would be a wrong prediction. Arregi et al. simply stipulate that this reading is not available. And they are in fact empirically hamstrung by the (once again overlooked) existence of reverse specificational. Given these are generally assumed (pace Hartmann, see 8.3.2) to be available, then there is no reason why this shouldn't be available as a reading of (64b), if focus is on *3.99*. But evidently the scope of negation will remain such that only the 'predicational' reading of (66b) is available.

I think a more natural explanation of these facts would be along the lines of the way we analysed exhaustivity 8.4.1. We can hold the semantics (in terms of semantic type, an $\langle e \rangle$ or $\langle e, t \rangle$ function) constant but explain the differences in terms of different scope relations, arising from the ordering of the flankers. The reason, which is very simple, that *one of the prices* scopes over negation in (64a) is that it c-commands *not* and the reason negation scopes over *one of the prices* in (64b) is that *there not* c-commands *one of the prices*.¹³

The other piece of evidence that Arregi et al. give is the possibility of a scalar reading with *only* in (67a), which is not possible in (67b):

¹³C-command may not in fact be a necessary underpinning of scope, see Safir (2004b), Barker (2012). But regardless, however scope is technically analysed, my contention is that the different scope relations arise from the different ordering of the copula's flankers rather than any putative semantic type difference in the analysis of the flankers.

- (67) a. One of the prices is only 3.99.
 b. Only 3.99 is one of the prices.

(67a) is true if and only if there is a price p such that p is 3.99 and no higher. (67b) cannot have this reading (it isn't entirely clear to me what reading it does have; it seems like it's saying that one of the prices only has one price, which is 3.99. But then prices of things generally are assumed to be the *only* prices of those things). Again, Arregi et al. claim that their analysis of 'specificationals' like (67a) capture the scalar reading, where the inverse predication reading fails. Specifically they show that the inverse predication analysis gives rise to the reading for (67a) that 3.99 is the highest price among the prices, which is not what the sentence says.

Here I think Arregi et al. give the answer to the problem themselves. They note that *only* only gets a scalar reading when it is attached to a VP, not an NP. They give the examples:

- (68) a. She only hired a janitor.
 b. She hired only a janitor.

The scalar reading where there is an inference that she didn't hire someone of higher status only goes with (68a), where *only* attaches to the VP *hired the janitor*. Where *only* attaches to the NP *a janitor* in (68b) this reading is not available (and here there is a clear reading where she hired a janitor and nobody else). Where copular sentences are concerned, I think it's fairly obvious that only (67a) can provide a reading where *only* can be understood as attaching to VP (in theory (67a) is ambiguous between the scalar reading and the implausible reading I suggested just above may be the only one possible in (67b)); as Arregi et al. say, finite forms of *be* surface to the immediate left of left-edge VP modifiers quite generally (*John is rarely amusing, Mary is often in the garden*). However, appearing to the left of the subject is not a position where *only* can be taken to be attached to the VP.¹⁴ And so the reading where *only* associates only with 3.99 in (67b) arises regardless of the semantics of *one of the prices*. And again, Arregi et al. would have to say why there is not an individual concept analysis for (67b) (given the generally accepted view that it can be a reverse specificational), and why—if they accepted that it could have such an analysis—we still wouldn't get the scalar reading of *only*, unlike what their analysis would predict we should get.

This section was merely intended to present some arguments from the very recent literature which explicitly claim that there is a truth-functional non-equivalence between (1) and (2)

¹⁴This is different to the other examples I just gave. We can say *Rarely, John is amusing, Often, Mary is in the garden*. But then 'rarely John' and 'often Mary' doesn't make sense (unlike 'Only John/Mary'), so there is no competing interpretation.

and that this is related to properly semantic distinctions between them to do with semantic type ambiguities of their flankers. I hope to have shown that the arguments pose no threat to the view I propose here, which holds that we keep the semantic types of the constituents constant: the scope difference related to negation and *only* arises independently of the semantic type analysis we apply to the flankers.

8.6 Conclusion

This chapter has tried to spell out a little more concretely the flanker-based, more pluralistic (in terms of the overall number of copular sentence types) approach that the preceding chapters lend themselves to. I've tried to keep this as simple as possible in order to prove the principle that this approach can work. For this reason I chose a minimal pair *John is the culprit*, *The culprit is John*. I wanted to show that we could account for their specificational vs. predicational property without resorting to any syntax–semantics distinctions; the distinctions we resorted to were (1) the way the flankers referred (attributively or referentially) and (2) which flanker was focused and which was the background. I scaled this up to a pseudocleft case in 8.4.4, which didn't seem too problematic, and then confronted it with Higgins's eight-ways ambiguous example that we met in the introduction 1.3, which was more problematic, though hopefully not fatally.

So the approach here ultimately gives rise to a description for any given sentence in terms of a flanker property statement. And this just is what I'm putting forward as a 'type of sentence' of a potential taxonomy of copular sentences. For the specificational *The culprit is John*, therefore, the description here is:

- (69) A syntactic predication where the pre-copula constituent is a background attributive DP and the post-copula constituent is an (exhaustively) focused referential DP.

and similarly for the entire taxonomy of copular sentences, depending on what the flanker properties are adjudged to be. They are all syntactic predications in this system; the differences lie in the way they refer and their information structure. In section 6.5, I admitted that I wasn't sure about what the relation between flanker properties and sentence functions is. Whilst this is still the case, with a description of a sentence as in (69), it certainly looks like there would be quite a transparent relation between them. If we called 'background attributive DP' a 'list heading' and 'focused referential DP' a 'list item', then specification boils down to a list heading and a list item i.e. the sentence function of listing, as Higgins put it.

In chapter section 6.5, I made the point that it seemed the copular taxonomy—understood

as a taxonomy of meanings relating (at least partly) to the way we use I-language—could only be described, not explained. However, having gone through this chapter the reader may feel that we really are explaining the taxonomy. The reason for considering that the taxonomy can be described, not explained, is the principled one that McGilvray gives in (Chomsky & McGilvray, 2012), namely that functions are uses of I-language i.e. they are actions, and ‘action is free’ i.e. governed by free will and hence not the province of scientific theory. However, in so far as we could in theory exhaustively list the various flanker properties (of which at least some were given in section 7.7.2), it isn’t obvious to me that we wouldn’t be doing some kind of explanation; an exhaustive specification of all possible flanker properties would indeed predict all possible permutations of flanker properties. Which of these are licit and illicit may also be rule-governed. At that point, it would certainly appear as though we have an explanation of the taxonomy of copular sentences, more so than a description.

Chapter 9

Conclusion

9.1 Summary

In many ways, chapter 7 was a ‘pre-conclusion’, collating the conclusions arrived at thus far in the thesis, dealing with some outstanding issues, and presenting more systematically the approach to describing the copular taxonomy that I think the findings of the previous chapters lead to. This then cleared the way for me to have a go at fleshing out the approach in a case study in chapter 8. Briefly, then, I’ll run through the thesis here.

The problem I started with was how to explain the taxonomy of copular sentences as laid out in Higgins (1979). Whilst primarily undertaken as a problem in linguistics, this is an interesting problem for philosophers too. Copular sentences (like any natural world phenomenon) can be taxonomically categorised in many different ways. For example, they can be categorised according to syntactic properties (such as the agreement patterns between the flankers and the copula, whether the flankers invert or not, and ‘connectivity effects’—see 7.5.5.2). Or if one’s interest is in the spatio-temporal properties of the flankers, one can divide copular sentences up according to this (as Roy (2013) does); one could even be interested in the metaphysical categories referred to by copular sentences (see 7.3) and classify them accordingly. Higgins’s taxonomy is quite clearly a taxonomy of the possible *meanings* of copular sentences, and as such it is philosophically interesting to try and understand where these distinct meanings come from, particularly as the meanings relate to the notions of predication and identity. In the broadest perspective, then, this thesis is an attempt to bring the combined forces of modern linguistics and philosophy together to understand how we make meaningfully distinct sentences. It was useful to focus on copular sentences for this task as a tool of exposition because they are the minimally-sized sentences in many languages. Another main motivation to exploring copular sentences was the debate in linguistics about the status of specificational sentences as predications or identities (e.g. Moro (1997), Heycock

and Kroch (1999)); it seemed to me that the debate was stuck and, perhaps presumptuously, I felt that it would help the debate to take a step back from the linguistic evidence for one side or the other, in order to see if there were any conceptual clarifications that might help break the impasse.

I started in what I took to be the obvious place—predicate logic, which contains predication and identity as primitives. If it is true that natural language copular sentences can be reduced to exclusively logical predications and logical identities, then there is perhaps hope for saying that logic explains the copular sentences we have. This was the work of chapters 2 and 3. The conclusions were (1) with respect to logical predication: it doesn't seem to be a truly semantic notion, but rather part of the combinatorics of predicate logic; as such it couldn't give rise to the notion of 'property ascription' that many linguists assume predication does give rise to (2) with respect to identity: (a) the notion of numerical identity from metaphysics does not feed into the meanings of informative sentences; (b) the study of the meanings of identity sentences in analytic philosophy (notably Frege's work) is more promising, but the most promising theory (Frege's 'On Sense and Reference' theory) only applies to identity statements i.e. equatives like *Hesperus is Phosphorus*, and it wouldn't obviously make sense to extend it to any of the other copular sentences, like specificationals or identificationals.

Chapters 4 and 5 didn't really advance the argument much; they were aimed at anyone (like Russell (1919, 172)) who is particularly tempted by the idea that the copula itself must be lexically ambiguous between the meanings of predication and identity. Chapter 4 just says that the copula isn't a lexical verb (or even really a proper verb in English) and chapter 5 tries to satisfy anyone dissatisfied with chapter 4 by showing some examples of why we do have copulas.

Chapter 6 goes back to the beginning to some extent to try and clarify what Higgins may have had in mind when he originally divided his taxonomy up into 'functions' a term about which he expressed 'conceptual difficulties'. Sentences can, I think, be given functions, along the lines of Scott-Phillips's (2015b) work, but Chomsky's views on this matter are significant and trying to elucidate them helped me get clearer on how to categorise the pieces of the puzzle that I think should enter into the description of copular sentences. I also noted that there is not obviously a substantial conflict between Chomsky's views on language and communication and those of mainstream pragmatics, even though I suspect many believe there is. This chapter also clarified that we cannot aim for an *explanation* of copular sentence meanings, but only a *description* of them, see section 6.5. Conceptual difficulties remain, in particular the relation between copular flanker properties and copular sentence functions; section 6.5.1 shows, with an example from Irish in Roy (2013, ch.7), that sentence function

cannot be fully determined by flanker properties. In the light of this, I suggested that we abandon the notion of a sentence function and deal solely in terms of flanker properties for the purposes of distinguishing meanings. But there may be an interesting research project related to examining the limits of the ‘inter-translatability’ of different flanker properties in different languages. As I said, chapter 7 collated my conclusions thus far and led up to a list of factors that I think enter into the description of the different meanings of copular sentences, according to a ‘full description’ which is a multifactorial model of flanker-property permutations. Information structure (understood as something distinct to syntax here), in particular, plays a major distinguishing role; we had already encountered a need for it in Frege’s work 2.2.3, in terms of deciding which function for a sentence like *Carbon dioxide is heavier than hydrogen* is in fact the main function of the sentence. Syntactic analyses like Roy (2013) are re-classified as a *use* of I-language to refer in different ways, see 7.5.5.3. Syntax plays a core foundational role in supplying the combinatorial capacity that allows us to put the flankers together in the first place (Merge, re-analysed as ‘phasal coincidence’, section 7.5.3); but it doesn’t distinguish the meanings of any sentences here. Chapter 8 tried to show this in action. The approach, overall, doesn’t take Higgins’s original taxonomy as a fact of the matter, and generally envisages a more pluralistic description of the meaning distinctions between copular sentences.

9.2 Desiderata

There have been a number of outstanding issues throughout the thesis. I just want to mention three briefly here.

9.2.1 A relation between the flankers? And what is predication?

As I say, this thesis has dealt in ‘flanker properties’ (such as the attributive vs. referential use of a flanker, or the focus vs. background use of a flanker) with a uniform and semantically vacuous notion of syntactic predication. Not specifying a relation between the flankers has been something of an application of Ockham’s razor; I haven’t needed relations of any sort in order to make meaning distinctions once the flanker properties are specified. But I’m uneasy with the notion that there is no meaningful relation of any sort between the flankers. My own hunch about this issue is that there is in fact some kind of abstract relation that could be described as a very basic kind of ‘modifying’ or ‘information updating’ (and which is uniform across copular sentences), possibly in line with work by Groenendijk and Stokhof (1991). It would be useful to have such a notion, as I think it would be necessary in

order to actually know what predication is. As I've noted in a few places, syntax gives us a necessary component of predication—the unit-making component, see 7.5—but doesn't obviously give us a sufficient condition (though work by Collins (2011) and Reichard (2013) is more optimistic about the role of syntax here, or a broader notion of 'grammar' which includes syntax in the case of Reichard.). It hasn't mattered too much for the aims of this thesis, as I say, as I have been assuming that copular sentences are all uniformly syntactic predications, so it wouldn't distinguish them anyway. But it would be nice to know.

9.2.2 What is identity?

I confess, too, that I haven't presented an analysis of equatives here either. Equatives are tricky beasts, as many who have undertaken to study them have discovered. Chapter 3 suggested that there wasn't necessarily only one type of equative, though Frege's 'On Sense and Reference' theory perhaps captures the 'classic' cases, and I said it might be promising to follow this up in a more modern linguistic idiom. Linguists have made the point that one of the flankers is not 'normally referring' by treating identity sentences as predications. For den Dikken (2006a), for example, the subject of an equative is contained within a null-headed predicate, such that *Cicero is Tully* could be paraphrased

- (1) (The one who is) Cicero is Tully.

This has a semantic counterpart in the type-shifting of one of the flankers to a predicative expression (using the IDENT type-shifting function of Partee (1986a), see sections 2.3.6 and 8.3.1) which yields 'the singleton set of individuals that are equal to *Cicero*' (den Dikken & O'Neill, 2017, 20). This, I think, can be couched in terms of the sort of tools that I used to deal with specificational and predicational distinctions in chapter 8. As den Dikken and O'Neill (2017, 20) put it:

[T]he two relata in an equative copular sentence are asymmetrical both in syntax and in information structure... [I]dentity statements... are useful in situations where an interlocutor is acquainted with an entity or individual under two different descriptions and unaware that the descriptions pick out the same individual, or acquainted with an entity but unaware that it is associated with two different descriptions or names. An approach to this asymmetry is to treat at least one of the expressions as a nonreferential or intensional object.

This is consistent with other recent work; for example Roy and Shlonsky (forthcoming) say:

The claim here was that the obligatory presence of the copula exactly corresponds with an equative reading. As I said in section 4.2.3 I don't precisely know how to analyse this evidence but I accept that evidence like this must be accounted for in other terms that don't resort to predication and identity as primitives. This would really be first-order linguistic work though, certainly requiring a greater amount of data from Hebrew in the case of (2). However, I am optimistic that such data can be explained without too much difficulty using the sorts of tools I used to distinguish predicational from specificational readings in chapter 8. This then would really help to clear the path to understanding meaning distinctions in copular sentences in more appropriate terms than logic and syntax-based analyses have accomplished thus far.

References

- Abney, Steven P. 1987. *The English noun phrase in its sentential aspect*. Ph.D. thesis, Massachusetts Institute of Technology.
- Adger, David, & Ramchand, Gillian. 2003. Predication and equation. *Linguistic Inquiry*, **34**(3), 325–59.
- Akmajian, Adrian. 1970. *Aspects of the grammar of focus in English*. Ph.D. thesis, MIT.
- Anselm of Canterbury. 2000. On Truth/De Veritate. *Pages 164–90 of: Complete philosophical and theological treatises of Anselm of Canterbury*. Minneapolis, MN: The Arthur J. Banning Press. Trans: J. Hopkins & H. Richardson.
- Arche, María J., Fábregas, Antonio, & Marín, Rafael. forthcoming. Main questions in the study of copulas: Categories, structures, and operations. In: Arche, María J., Fábregas, Antonio, & Marín, Rafael (eds), *The grammar of copulas across languages*. Oxford: Oxford University Press.
- Aristotle. 1963. *Categories and de interpretatione*. Oxford: Clarendon Press. Trans: Ackrill, J.L.
- Arregi, Karlos, Francez, Itamar, & Martinović, Martina. 2018. *Specificational subjects are individual concepts*. MS, University of Chicago, <http://ling.auf.net/lingbuzz/003996>.
- Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy*, **9**, 5–16.
- Baddeley, Alan. 2007. *Working memory, thought, and action*. Oxford: Oxford University Press.
- Bahloul, Maher. 2006. Copula. *Pages 506–11 of: Versteegh, Kees, Eid, Mushira, Elgibali, Alaa, Woidich, Manfred, & Zaborski, Andrzej (eds), Encyclopedia of Arabic Language and Linguistics*, vol. 1. Leiden: Brill.
- Baker, C. L. 1968. *Indirect questions in English*. Ph.D. thesis, University of Illinois.

- Baker, Mark C. 2001. *Atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.
- Barker, Chris. 2012. Quantificational binding does not require c-command. *Linguistic inquiry*, **43**(4), 614–633.
- Barros, Matthew. 2016. *Specification as an inter-taxonomic phenomenon*. MS, Yale University, <http://ling.auf.net/lingbuzz/003117>.
- Basilico, David. 2003. The topic of small clauses. *Linguistic Inquiry*, **34**(1), 1–35.
- Becker, Misha Karen. 2000. *The development of the copula in child English: The lightness of Be*. Ph.D. thesis, University of Pennsylvania.
- Bejar, Susana, Denniss, Jessica, Kahnemuyipour, Arsalan, & Yokoyama, Tomohiro. forthcoming. Number matching in binomial small clauses. In: Arche, María J., Fábregas, Antonio, & Marín, Rafael (eds), *The grammar of copulas across languages*. Oxford: Oxford University Press.
- Belletti, Adriana. 2001. Agreement projections. In: Baltin, M., & Collins, C. (eds), *The handbook of contemporary syntactic theory*. Oxford: Blackwell.
- Benveniste, Émile. 1966. La phrase nominale. *Pages 151–167 of: Problèmes de linguistique générale*. Paris: Gallimard.
- Benz, Anton, & Jasinskaja, Katja. 2017. Questions Under Discussion: From sentence to discourse. *Discourse Processes*, **54**(3), 177–186.
- Beyssade, Claire, & Dobrovie-Sorin, Carmen. 2009 (March). *Copular sentences and identity*. Handout, Tucson, Arizona.
- Biberauer, Theresa. 2004. Reconsidering the EPP and Spec-TP in Germanic. *Cambridge Occasional Papers in Linguistics*, **1**, 15–39. Lluïsa Astruc & Marc Richards (eds).
- Biberauer, Theresa, & Richards, Marc. 2006. True Optionality: When the grammar doesn't mind. In: Boeckx, Cedric (ed), *Minimalist Essays*. Amsterdam: John Benjamins.
- Birner, Betty J. 1994. Information status and word order: An analysis of English inversion. *Language*, 233–59.
- Birner, Betty J. 1996. *The discourse function of inversion in English*. New York: Garland.

- Birner, Betty J, Kaplan, Jeffrey P, & Ward, Gregory L. 2007. Functional compositionality and the interaction of discourse constraints. *Language*, **83**(2), 317–343.
- Black, Max. 1952. The identity of indiscernibles. *Mind*, **61**(242), 153–164.
- Blanchette, Patricia. 2002. Review of *Logical properties* by C. McGinn. *Notre Dame Philosophical Reviews*, March. Available at: <https://ndpr.nd.edu/news/logical-properties/>.
- Blom, Alied, & Daalder, Saskia. 1977. *Syntaktische theorie en taalbeschrijving*. Muiderberg: Coutinho.
- Boeckx, Cedric. 2014. *Elementary syntactic structures: Prospects of a feature-free Syntax*. Cambridge: Cambridge University Press.
- Borer, Hagit. 2005a. *In name only. Structuring sense: Volume 1*. Oxford: Oxford University Press.
- Borer, Hagit. 2005b. *The normal course of events. Structuring sense: Volume 2*. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Taking form. Structuring sense: Volume 3*. Oxford: Oxford University Press.
- Bowers, John. 1991. The syntax and semantics of nominals. *Pages 1–30 of: Cornell Working Papers in Linguistics*, vol. 10. Ithaca, NY: Cornell University.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry*, **24**(4), 591–656.
- Bowers, John. 2001. Predication. In: Baltin, M., & Collins, C. (eds), *The handbook of contemporary syntactic theory*. Oxford: Blackwell.
- Brody, Michael. 1995. *Lexico-Logical Form: A radically minimalist theory*. Cambridge, MA: MIT press.
- Brody, Michael. 2002. On the status of derivations and representations. In: Epstein, S., & Seely, D. (eds), *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell.
- Carlson, Gregory. 1977. *Reference to kinds in English*. Ph.D. thesis, University of Massachusetts.
- Carnie, Andrew. 1995. *Non-verbal predication and head-movement*. Ph.D. thesis, MIT.

- Chierchia, Gennaro. 1984. *Topics in the syntax and semantics of infinitives and gerunds*. Ph.D. thesis, University of Massachusetts, Amherst.
- Chierchia, Gennaro. 1985. Formal semantics and the grammar of predication. *Linguistic Inquiry*, **16**(3), 417–443.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1970. Remarks on nominalization. *Pages 184–221 of: Jacobs, R., & Rosenbaum, P. (eds), Readings in English transformational grammar*. Waltham, MA: Blaisdell Publishing.
- Chomsky, Noam. 1972. *Language and mind*. New York: Harcourt Brace Jovanovich.
- Chomsky, Noam. 1975. *The logical structure of linguistic theory*. New York: Plenum Press.
- Chomsky, Noam. 1980. *Rules and representations*. Oxford: Basil Blackwell.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1983. Some conceptual shifts in the study of language. *In: L. Cauman, I. Levi, C. Parsons, & Schwartz, R. (eds), How Many Questions? Essays in honor of Sidney Morgenbesser*. Indianapolis, IN: Hackett.
- Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 1988. *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: MIT press.
- Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. *Pages 89–155 of: Martin, Roger, Michaels, David, & Uriagereka, Juan (eds), Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. *Pages 1–52 of: Kenstowicz, Michael (ed), Ken Hale: A Life in Language*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. *Pages 104–31 of: Belletti, Adriana (ed), Structures and beyond*. Oxford: Oxford University Press.

- Chomsky, Noam. 2008. On phases. *Pages 133–66 of: Freidin, Robert, & Otero, Carlos P. (eds), Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud.* Cambridge, MA: MIT Press.
- Chomsky, Noam. 2009. *Noam Chomsky and Samuel Jay Keyser.* Podcast, Cambridge, MA. Available at: <https://mitpress.podbean.com/e/noam-chomsky-and-samuel-jay-keyser/>.
- Chomsky, Noam. 2013. Problems of projection. *Lingua*, **130**, 33–49.
- Chomsky, Noam. 2015. Problems of projection: Extensions. *In: Domenico, Elisa Di, Hamann, Cornelia, & Matteini, Simona (eds), Structures, strategies and beyond: Studies in honour of Adriana Belletti.* Amsterdam/Philadelphia: John Benjamins.
- Chomsky, Noam. 2016 (September). *Noam Chomsky on the Evolution of Language: A Bilingual Perspective.* <https://truthout.org/articles/noam-chomsky-on-the-evolution-of-language-a-bilingual-perspective/>. Last accessed: 9th July 2018.
- Chomsky, Noam, & Lasnik, Howard. 1993. The theory of principles and parameters. *Pages 506–569 of: Jacobs, Joachim, von Stechow, Arnim, Sternefeld, Wolfgang, & Vennemann, Theo (eds), Syntax: An international handbook of contemporary research*, vol. 1. Berlin: Walter de Gruyter.
- Chomsky, Noam, & McGilvray, James. 2012. *The science of language: Interviews with James McGilvray.* Cambridge: Cambridge University Press.
- Cinque, G. 2010. *The Syntax of adjectives: A comparative study.* Cambridge, MA: MIT Press.
- Cinque, G. 2014. Again on Tense, Aspect, Mood morpheme order and the “Mirror Principle”. *In: Svenonius, Peter (ed), Functional Structure from Top to Toe.* Oxford: Oxford University Press.
- Citko, Barbara. 2014. *Phase theory: An introduction.* Cambridge: Cambridge University Press.
- Cohen, S. Marc. 2008. *Philosophy 453: Philosophy of Language lecture notes.* University of Washington. Available at: <https://faculty.washington.edu/smcohen/453/LogicHandout.pdf>.
- Cohen, S. Marc. 2016. Aristotle’s Metaphysics. *In: Zalta, Edward N. (ed), The Stanford Encyclopedia of Philosophy*, winter 2016 edn. Metaphysics Research Lab, Stanford University.

- Collins, John. 2011. *The unity of linguistic meaning*. Oxford: Oxford University Press.
- Comorovski, Ileana. 2007. Constituent questions and the copula of specification. In: Comorovski, Ileana, & von Stechow, Klaus (eds), *Existence: Semantics and syntax*. Studies in Linguistics and Philosophy, vol. 84. Dordrecht: Springer.
- Cottingham, John (ed). 2008. *Western philosophy: An anthology, 2nd edn*. Oxford: Blackwell.
- Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.
- Cumming, Samuel. 2008. Variabilism. *Philosophical Review*, **117**(4), 525–54.
- Darwin, Charles. 1859. *The origin of species*. Oxford: Oxford University Press, 2008 edition.
- Davidson, Donald. 2001. *Essays on actions and events*. Oxford: Oxford University Press.
- Davidson, Donald. 2005. *Truth and predication*. Cambridge, MA: Harvard University Press.
- Declerck, Renaat. 1988. *Studies on copular sentences, clefts, and pseudo-clefts*. Dordrecht: Leuven University Press.
- den Dikken, Marcel. 2006a. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press.
- den Dikken, Marcel. 2006b. Specificational copular sentences and pseudoclefts. In: Everaert, Martin, & van Riemsdijk, Henk (eds), *The Blackwell companion to syntax*. Oxford: Wiley-Blackwell.
- den Dikken, Marcel. 2008 (27 March). *Small clauses, phases, and phase extension—the case of scope and object shift*. Paper presented 31st GLOW Colloquium. https://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Programs/Linguistics/Dikken/scope_paper_glow.pdf.
- den Dikken, Marcel, & O'Neill, Teresa. 2017. Copular constructions in syntax. In: Aronoff, Mark (ed), *Oxford research encyclopedia of linguistics*. Oxford University Press. <http://oxfordindex.oup.com/view/10.1093/acrefore/9780199384655.013.137>.
- Dickie, Imogen. 2008. Informative identities in the Begriffsschrift and 'On Sense and Reference'. *Canadian Journal of Philosophy*, **38**(2), 269–287.

- Dik, Simon C. 1983. Auxiliary and copula be in a functional grammar of English. *Pages 121–143 of: Heny, Frank, & Richards, Barry (eds), Linguistic categories: auxiliaries and related puzzles.* Dordrecht: Reidel.
- Dik, Simon C. 1989. *The theory of functional grammar.* Dordrecht: Foris.
- Donnellan, Keith S. 1966. Reference and definite descriptions. *The philosophical review*, **75**(3), 281–304.
- Doron, E. 1983. *Verbless predicates in Hebrew.* Ph.D. thesis, University of Texas at Austin.
- Doucet, Sébastien, Soussignan, Robert, Sagot, Paul, & Schaal, Benoist. 2009. The secretion of areolar (Montgomery’s) glands from lactating women elicits selective, unconditional responses in neonates. *PLoS One*, **4**(10), e7579.
- Dummett, Michael. 1981. *Frege: Philosophy of language, 2nd edn.* London: Duckworth.
- Emonds, J. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations.* New York: Academic Press.
- Englebretsen, George. 1980. A note on predication. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, **19**(4), 627–628.
- Englebretsen, George, & Sayward, Charles. 2011. *Philosophical logic: An introduction to advanced topics.* London and New York: Continuum.
- Faarlund, Jan T. 2007. From cleft sentence to morphological focus marker in Zoque. *Pages 221–30 of: Nilsen, Randi A., Amfo, Nana Aba Appiah, & Borthen, Kaja (eds), Interpreting utterances: Pragmatics and its interfaces. Essays in honour of Thorstein Fretheim.* Oslo: Novus.
- Ferguson, Charles A. 1971. Absence of copula and the notion of simplicity. *Pages 141–150 of: Hymes, Dell (ed), Pidginization and creolization of languages.* Cambridge: Cambridge University Press.
- Féry, Caroline, & Krifka, Manfred. 2008. Information structure: Notional distinctions, ways of expression. *In: van Sterkenburg, Piet (ed), Unity and diversity of languages.* John Benjamins Publishing.
- Forrest, Peter. 2016. The Identity of Indiscernibles. *In: Zalta, Edward N. (ed), The Stanford Encyclopedia of Philosophy*, winter 2016 edn. Metaphysics Research Lab, Stanford University.

- Frege, Gottlob. 1879. Begriffsschrift. *Pages 1–21 of*: Geach, P., & Black, M. (eds), *Translations from the philosophical writings of Gottlob Frege, 2nd revised ed., 1970*. Oxford: Blackwell.
- Frege, Gottlob. 1884. *The foundations of arithmetic: A logico-mathematical enquiry into the concept of number*. Evanston, IL: Northwestern University Press, 1980.
- Frege, Gottlob. 1892a. On concept and object. *Pages 42–56 of*: Geach, P., & Black, M. (eds), *Translations from the philosophical writings of Gottlob Frege, 2nd revised ed., 1970*. Oxford: Blackwell.
- Frege, Gottlob. 1892b. On sense and reference. *Pages 56–78 of*: Geach, P., & Black, M. (eds), *Translations from the philosophical writings of Gottlob Frege, 2nd revised ed., 1970*. Oxford: Blackwell.
- Frege, Gottlob. 1893/1903. *Grundgesetze der Arithmetik (Basic laws of arithmetic: Derived using concept-script)*. Oxford: Oxford University Press, 2013. Trans: P. Ebert and M. Rossberg (with C. Wright).
- Gallego, Ángel, & Uriagereka, Juan. 2016. $Estar=Ser+X$. *Borealis—An International Journal of Hispanic Linguistics*, 5(1), 123–56.
- Gallego, Ángel. 2010. *Phase theory*. Amsterdam: John Benjamins.
- Gamut, L. T. F. 1991a. *Logic, language, and meaning: Intensional logic and logical grammar*. Vol. 2. Chicago: University of Chicago Press.
- Gamut, L. T. F. 1991b. *Logic, language, and meaning: Introduction to logic*. Vol. 1. Chicago: University of Chicago Press.
- Geach, P. T. 1968. *Reference and generality*. Ithaca, NY: Cornell University Press. Emended edition.
- Geach, P. T. 1972. *Logic matters*. Oxford: Basil Blackwell.
- Geach, P. T. 1980a. *Reference and generality*. 3rd edn. London: Cornell University Press.
- Geach, P. T. 1980b. Strawson on Subject and Predicate. *Pages 174–188 of*: Straaten, Z. Van (ed), *Philosophical Subjects*. Oxford University Press.
- Geist, Ljudmila. 2007. Constituent questions and the copula of specification. *In*: Comorovski, Ileana, & von Heusinger, Klaus (eds), *Existence: Semantics and Syntax*. Studies in Linguistics and Philosophy, vol. 84. Dordrecht: Springer.

- Gelstein, Shani, Yeshurun, Yaara, Rozenkrantz, Liron, Shushan, Sagit, Frumin, Idan, Roth, Yehudah, & Sobel, Noam. 2011. Human tears contain a chemosignal. *Science*, **331**(6014), 226–230.
- Gödel, Kurt. 1990. Russell's mathematical logic. *Pages 119–44 of*: Feferman, S., Dawson Jr, J.W., Kleene, S.C., Moore, G.H., Solovay, R.M., & van Heijenoort, J. (eds), *Kurt Gödel: Collected works*, vol. II. New York and Oxford: Oxford University Press.
- Grice, H. P. 1957. Meaning. *The philosophical review*, **66**(3), 377–388.
- Grice, H. P. 1975. Logic and conversation. *Pages 41–58 of*: Cole, P., & Morgan, J. (eds), *Syntax and semantics volume 3: Speech acts*. Academic Press.
- Grice, H. P. 1989. *Studies in the way of words*. Cambridge, MA: Harvard University Press.
- Groenendijk, Jeroen, & Stokhof, Martin. 1991. Dynamic predicate logic. *Linguistics and Philosophy*, **14**, 39–100.
- Haegeman, Liliane, & Guéron, Jacqueline. 1999. *English grammar: A generative perspective*. Oxford: Blackwell Publishing.
- Hale, K., & Keyser, S.J. 1993. On argument structure and the lexical expression of syntactic relations. *Pages 53–109 of*: Hale, K., & Keyser, S.J. (eds), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.
- Halle, Morris, & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. *Pages 111–76 of*: Hale, Kenneth, & Keyser, S. Jay (eds), *The View from Building 20*. Cambridge, MA: MIT Press.
- Halle, Morris, & Marantz, Alec. 1994. Some key features of Distributed Morphology. *Pages 275–88 of*: Carnie, Andrew, & Harley, Heidi (eds), *MIT working papers in linguistics: Papers on phonology and morphology*, vol. 21.
- Halliday, M. A. K. 1967. Notes on transitivity and theme in English. Parts I and II. *Journal of Linguistics*, **3**(1), 37–81, 199–244.
- Halliday, M. A. K., & Matthiessen, Christian. 2004. *An introduction to functional grammar*. London: Hodder Arnold.
- Hartmann, Jutta. forthcoming. Focus and Prosody in Nominal Copular Clauses. *In*: Featherston, S., Hörnig, R., von Wietersheim, S., & Winkler, S. (eds), *Information Structure and Semantic Processing*. Linguistische Arbeiten. Berlin: De Gruyter.

- Hartmann, Jutta, & Hegedűs, Veronika. 2009. *Nominal copula clauses in English and Hungarian*. manuscript.
- Hartmann, Katharina, & Veenstra, Tonjes. 2013. Introduction. In: Hartmann, Katharina, & Veenstra, Tonjes (eds), *Cleft structures*. Amsterdam/Philadelphia: John Benjamins Publishing.
- Hauser, Marc D., Chomsky, Noam, & Fitch, W. Tecumseh. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science*, **298**(5598), 1569–1579.
- Heggie, Lorie. 1988. *The syntax of copular structures*. Ph.D. thesis, University of Southern California.
- Heim, Irene, & Kratzer, Angelika. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Heller, Daphna. 2005. *Identity and information: Semantic and pragmatic aspects of identificational sentences*. Ph.D. thesis, Rutgers University.
- Heller, Daphna, & Wolter, Lynsey. 2008. ‘That is Rosa’: identificational sentences as intensional predication. In: Gronn, Atle (ed), *Proceedings of SuB12*. Oslo: ILOS.
- Heycock, Caroline. 1991. *Layers of predication*. Ph.D. thesis, University of Pennsylvania.
- Heycock, Caroline. 1992. Layers of predication and the syntax of the copula. *Belgian Journal of Linguistics*, **7**, 95–123.
- Heycock, Caroline. 1994. *Layers of predication*. New York: Garland.
- Heycock, Caroline. 2012. Specification, equation, and agreement in copular sentences. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, **57**(2), 209–240.
- Heycock, Caroline, & Kroch, Anthony. 1999. Pseudocleft connectivity: Implications for the LF interface. *Linguistic Inquiry*, **30**(3).
- Hicks, Glyn. 2006. *The derivation of anaphoric relations*. Ph.D. thesis, University of York.
- Higginbotham, J. 1985. On semantics. *Linguistic Inquiry*, **16**, 547–93.
- Higgins, F. R. 1979. *The pseudo-cleft construction in English*. New York: Garland.
- Hinzen, Wolfram. 2006. *Mind design and minimal syntax*. Oxford: Oxford University Press.

- Hinzen, Wolfram. 2012a. Minimalism. *Pages 93 – 141 of:* Kempson, Ruth, Fernando, Tim, & Asher, Nicholas (eds), *Philosophy of Linguistics*. Handbook of the Philosophy of Science. Amsterdam: North-Holland.
- Hinzen, Wolfram. 2012b. The philosophical significance of Universal Grammar. *Language Sciences*, **34**(5), 635–649.
- Hinzen, Wolfram. 2016. On the grammar of referential dependence. *Studies in Logic, Grammar, and Rhetoric*, **46**(1), 11–33.
- Hinzen, Wolfram, & Sheehan, Michelle. 2014. *The philosophy of universal grammar*. Oxford: Oxford University Press.
- Horn, Laurence. 2016. Information structure and the landscape of (non-) at-issue meaning. *In:* Féry, Caroline, & Ishihara, Shinichiro (eds), *The Oxford Handbook of Information Structure*. Oxford: Oxford University Press.
- Horn, Laurence R. 1989. *A natural history of negation*.
- Hornstein, Norbert. 1995. *Logical form: From GB to minimalism*. Oxford: Blackwell.
- Hudson, Richard. 2003. Gerunds without phrase structure. *Natural Language & Linguistic Theory*, **21**(3), 579–615.
- Hume, David. 1748. *An enquiry concerning human understanding*. Oxford: Clarendon, 2000. Beauchamp, Tom L. (ed).
- Jackendoff, Ray. 1977. *X-bar syntax*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- Jespersen, Otto. 1924. *Philosophy of grammar*. London: Allen & Unwin.
- Jespersen, Otto. 1933. *The system of grammar*. London: Allen & Unwin.
- Kaplan, David. 1989. Demonstratives. *Pages 481–563 of:* Joseph Almog, John Perry, & Wettstein, Howard (eds), *Themes from Kaplan*. Oxford: Oxford University Press.
- Kayne, Richard S. 1984. *Connectedness and binary branching*. Dordrecht: Foris Publications.
- Kayne, Richard S. 1994. *Antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Kibrik, Andrej A. 2011. *Reference in discourse*. Oxford: Oxford University Press.

- Kratzer, Angelika. 1996. Severing the external argument from its verb. *Pages 109–137 of:* Rooryck, J., & Zaring, L. (eds), *Phrase structure and the lexicon*. Dordrecht: Springer.
- Kretzmann, N. 1982. Syncategoremata, sophismata, exponibilia. *Pages 211–45 of:* Kretzmann, N., Kenny, A., & Pinborg, J. (eds), *Cambridge history of later medieval philosophy*. Cambridge: Cambridge University Press.
- Krifka, Manfred, & Musan, Renate. 2012. Information structure: Overview and linguistic issues. *Pages 1–44 of:* Krifka, Manfred, & Musan, Renate (eds), *The expression of information structure*. Berlin: Mouton de Gruyter.
- Kripke, Saul. 1972. Naming and necessity. *Pages 253–355 of:* Davidson, D., & Harman, G. (eds), *Semantics of natural language*. Dordrecht: Reidel.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Langacker, Ronald W. 1969. On pronominalization and the chain of command. *Pages 160–186 of:* Reibel, D.A., & Schane, S.A. (eds), *Modern studies in English: Readings in transformational grammar*. Englewood Cliffs, NJ: Prentice-Hall.
- Langacker, Ronald W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites*. Vol. 1. Stanford, CA: Stanford University Press.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry*, **19**(3), 335–391.
- Leibniz, Gottfried Wilhelm. 1840. *Opera philosophica omnia*. Berlin: Joannes Eduardus Erdmann. <https://archive.org/details/operaphilosophi00erdmgoog>.
- Lewis, David. 1975. Languages and Language. *Pages 3–35 of:* Gunderson, Keith (ed), *Minnesota Studies in the Philosophy of Science*. University of Minnesota Press.
- Linsky, Leonard. 1963. Reference and referents. *Pages 74–89 of:* Caton, C. E. (ed), *Philosophy and ordinary language*. Urbana: University of Illinois Press.
- Lowe, E. J. 1980. Sortal terms and natural laws. *American Philosophical Quarterly*, **19**(4), 253–60.
- Lowe, E. J. 1982. Laws, dispositions and sortal logic. *American Philosophical Quarterly*, **19**(1), 41–50.
- Lowe, E. J. 2002. *A survey of metaphysics*. Oxford: Oxford University Press.

- Lowe, E. J. 2006. Review of Davidson, *Truth and Predication*. *Times Higher Education*, April 7th. <https://www.timeshighereducation.com/books/take-a-seat-then-consider-this-simple-sentence/202436.article>.
- Lowe, E. J. 2008. Tropes and perception. *Pages 175–192 of: Gozzano, S., & Orilia, F. (eds), Tropes, universals and the philosophy of mind: Essays at the boundary of ontology and philosophical psychology*. Frankfurt: Ontos Verlag.
- Lowe, E. J. 2009. *More Kinds of Being*. Chichester, West Sussex: Wiley-Blackwell.
- Lowe, E. J. 2013. *Forms of thought: A study in philosophical logic*. Cambridge University Press.
- Ludlow, Peter. 2018. Descriptions. *In: Zalta, Edward N. (ed), The Stanford Encyclopedia of Philosophy*, summer 2018 edn. Metaphysics Research Lab, Stanford University.
- Marelj, Marijana, & Matushansky, Ora. 2015. Mistaking *for*: Testing the theory of mediated predication. *Linguistic Inquiry*, **46**(1), 43–76.
- Martin, Txuss, & Hinzen, Wolfram. 2014. The grammar of the essential indexical. *Lingua*, **148**, 95–117.
- Martin, Txuss, Reichard, Ulrich, & Woodard, Andrew. 2016. *The topology of reference*. MS, Durham University.
- Matushansky, Ora. 2015. Against PredP. *In: Proceedings of IATL 2015*. <http://www.iatl.org.il/wp-content/uploads/2015/10/IATL30proceedings-06-matushansky.pdf>.
- Matushansky, Ora. to appear. Against the PredP theory of small clauses. *Linguistic Inquiry*.
- McGinn, Colin. 2000. *Logical properties*. Oxford: Oxford University Press.
- McGinn, Colin. 2015. *Philosophy of language: The classics explained*. Cambridge, MA: MIT Press.
- Meillet, Antoine. 1906. La phrase nominale en indo-européen. *Mémoires de la Société Linguistique de Paris*, **14**, 1–26.
- Mikkelsen, Line. 2005. *Copular clauses: Specification, predication, and equation*. Amsterdam: John Benjamins.

- Mikkelsen, Line. 2011. Copular clauses. *In: von Heusinger, Klaus, Maienborn, Claudia, & Pornter, Paul (eds), Semantics: An International Handbook of Natural Language Meaning*. HSK, vol. 33.2. Berlin: de Gruyter Mouton.
- Mill, John Stuart. 1856. *A system of logic*. Honolulu: University Press of the Pacific, 2002.
- Millikan, Ruth Garrett. 1984. *Language, thought, and other biological categories: New foundations for realism*. Cambridge, MA: MIT press.
- Milway, Daniel. 2016. *On some subjects of specificational copular clauses*. MS, University of Toronto, <http://ling.auf.net/lingbuzz/003058>.
- Moltmann, Friederike. 2013. Identificational sentences. *Natural Language Semantics*, **21**(1), 43–77.
- Moore, G. E. 1910–11. *Some main problems of philosophy*. London: George Allen & Unwin, 1953.
- Moro, Andrea. 1988. Per una teoria unificata delle frasi copulari. *Rivista di Grammatica Generativa*, 81–110.
- Moro, Andrea. 1991. The raising of predicates: copula, expletives and existence. *MIT working papers in linguistics*, **15**, 119–181.
- Moro, Andrea. 1997. *The raising of predicates*. Cambridge: Cambridge University Press.
- Moro, Andrea. 2006. Copular sentences. *In: Everaert, Martin, & Van Riemsdijk, Henk (eds), The Blackwell Companion to Syntax*. Oxford: Wiley-Blackwell.
- Moro, Andrea. 2013. *The equilibrium of human syntax: Symmetries in the brain*. New York: Routledge.
- Moulton, Keir. 2013. Small clause subjects do reconstruct. *Pages 27–40 of: Fainleib, Yelena, LaCara, Nicholas, & Park, Yangsook (eds), Proceedings of the Forty-First Annual North East Linguistic Society*. Amherst, MA: GLSA.
- Parascandola, Mark. 2011. Causes, risks, and probabilities: Probabilistic concepts of causation in chronic disease epidemiology. *Preventive Medicine*, **53**(4), 232 – 234. Special Section: Epidemiology, Risk, and Causation.
- Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

- Partee, Barbara. 1986a. Ambiguous pseudoclefts with unambiguous BE. *Pages 354–366 of: Berman, S., Choe, J. W., & McDonough, J. (eds), Proceedings of NELS 16.* University of Massachusetts, Amherst: GLSA.
- Partee, Barbara. 1986b. Noun phrase interpretation and type-shifting principles. *Pages 203–30 of: Compositionality in formal semantics: Selected papers by Barbara H. Partee, published 2004.* Oxford: Blackwell.
- Partee, Barbara, & Rooth, Mats. 1983. Generalized conjunction and type ambiguity. *Pages 334–56 of: Portner, P., & Partee, B. H. (eds), Formal semantics: the essential readings.* Oxford: Blackwell, 2002.
- Patten, Amanda. 2012. *The English It-Cleft: A constructional account and a diachronic investigation (Topics in English Linguistics).* Berlin/Boston: Mouton de Gruyter.
- Patten, Amanda. 2016. Well-formed lists: Specificational copular sentences as predicative inversion constructions. *English Language & Linguistics*, **22**(1), 77–99.
- Percus, Orin, & Sharvit, Yael. 2014. *Copular asymmetries in belief reports.* Paper presented the SALT 24. New York University.
- Pereltsvaig, Asya. 2001. *On the nature of intra-clausal relations.* Ph.D. thesis, McGill University, Montréal.
- Pinker, Steven. 1994. *The language instinct.* London: Penguin.
- Pinker, Steven, & Bloom, Paul. 1990. Natural language and natural selection. *Behavioral and Brain Sciences*, **13**(4), 707–727.
- Plato. c.380BC/1892. Republic [Politeia]. *Pages 507b1–517c6/pp.207–17 of: The dialogues of Plato*, vol. III. Oxford: Clarendon. Trans: B. Jowett.
- Pustet, Regina. 2003. *Copulas. Universals in the categorization of the lexicon.* Oxford: Oxford University Press.
- Quine, W. V. O. 1940. *Mathematical logic.* revised edition, 2009 edn. Cambridge, MA: Harvard University Press.
- Quine, W. V. O. 1943. Notes on existence and necessity. *The Journal of Philosophy*, **40**(5), 113–27.
- Quine, W. V. O. 1951. Two dogmas of empiricism. *The Philosophical Review*, **60**(1), 20–43.

- Quine, W. V. O. 1960. *Word and object*. Cambridge, MA: MIT Press.
- Quine, W. V. O. 1964. On what there is. *Pages 1–19 of: From a logical point of view*, 2nd revised edn. Cambridge, MA: Harvard University Press.
- Quine, W. V. O. 1981. Things and their places in theories. *Pages 1–23 of: Theories and things*. Cambridge, MA: Harvard University Press.
- Quirk, Randolph, & Greenbaum, Sidney. 1973. *A university grammar of English*. Harlow, Essex: Longman.
- Rapoport, Tova. 1985. Copular constructions in Hebrew. *In: Eilfort, W. H., Kroeber, P. D., & Peterson, K. L. (eds), CLS 21*. Chicago Linguistic Society.
- Rapoport, Tova. 1987. *Copular, nominal, and small clauses: A study of Israeli Hebrew*. Ph.D. thesis, MIT.
- Reichard, Ulrich. 2012. Making events redundant: Adnominal modification and phases. *In: Stalmaszczyk, Piotr (ed), Philosophical and formal approaches to linguistic analysis*. Frankfurt a. M.: Ontos.
- Reichard, Ulrich. 2013. *Grammar, ontology, and the unity of meaning*. Ph.D. thesis, Durham University, Durham, UK.
- Reichard, Ulrich, & Hinzen, Wolfram. 2016. Are verb names of events? A review of the event-argument hypothesis. *Pages 43–65 of: Stalmaszczyk, Piotr (ed), Philosophical and Linguistic Analyses of Reference (Studies in Philosophy of Language and Linguistics, vol. 2)*. Frankfurt: Peter Lang.
- Reinhart, Tanya. 1979. Syntactic domains for semantic rules. *Pages 107–130 of: Guenther, F., & Schmidt, S.J. (eds), Formal semantics and pragmatics for natural languages*. Dordrecht: Reidel.
- Reinhart, Tanya. 1982. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica*, **27**, 53–94.
- Ringe, Donald. 2006. *A linguistic history of English: Volume I, From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic inquiry*, **17**(3), 501–557.

- Romero, Maribel. 2005. Concealed questions and specificational subjects. *Linguistics and Philosophy*, **28**(6), 687–737.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics*, **1**(1), 75–116.
- Rothstein, Susan. 1983. *The syntactic forms of predication*. Ph.D. thesis, MIT.
- Rothstein, Susan. 2001. *Predicates and their subjects*. Dordrecht: Kluwer.
- Rothstein, Susan. 2006. Predication. *Pages 73–6 of: Brown, Keith (ed), Encyclopedia of language & linguistics, second edition*, vol. 10. Oxford: Elsevier.
- Roussarie, Laurent. 2006. Prédicat. In: Godard, Danièle, Roussarie, Laurent, & Corblin, Francis (eds), *Sémantique: dictionnaire de sémantique*. <http://www.semantique-gdr.net/dico/>: GDR Sémantique & Modélisation, CNRS.
- Rouveret, Alain, & Vergnaud, Jean-Roger. 1980. Specifying reference to the subject: French causatives and conditions on representations. *Linguistic Inquiry*, **11**(1), 97–202.
- Roy, Isabelle. 2013. *Nonverbal predication: Copular sentences at the syntax-semantics interface*. Oxford: Oxford University Press.
- Roy, Isabelle, & Shlonsky, Ur. forthcoming. Aspects of the syntax of ce in French copular sentences. In: Arche, María J., Fábregas, Antonio, & Marín, Rafael (eds), *The grammar of copulas across languages*. Oxford: Oxford University Press.
- Russell, Bertrand. 1903. *The principles of mathematics*. London: Allen and Unwin. 2nd edition, 1937.
- Russell, Bertrand. 1905. On denoting. *Mind*, **14**(56), 479–493.
- Russell, Bertrand. 1908. Mathematical logic as based on the theory of types. *American Journal of Mathematics*, **30**, 222–62.
- Russell, Bertrand. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press, 2001.
- Russell, Bertrand. 1917. *Mysticism and logic: and other essays*. London: Allen and Unwin.
- Russell, Bertrand. 1919. *The philosophy of mathematics*. London: Allen and Unwin.

- Russell, Bertrand. 1959. *My philosophical development, reprinted 1995 edition*. London: Routledge.
- Safir, Kenneth J. 1985. *Syntactic Chains*. Cambridge: Cambridge University Press.
- Safir, Kenneth J. 2004a. *The syntax of anaphora*. Oxford: Oxford University Press.
- Safir, Kenneth J. 2004b. *The syntax of (in)dependence*. Cambridge, MA: MIT Press.
- Safir, Kenneth J. 2005. Abandoning Coreference. In: Bermúdez, José Luis (ed), *Thought, Reference, and Experience: Themes from the Philosophy of Gareth Evans*. Oxford: Clarendon Press.
- Sandmann, Manfred. 1979. *Subject and predicate: a contribution to the theory of syntax*. Heidelberg: Winter.
- Scott-Phillips, Thom. 2015a (July). ‘Speaking our minds’ online book club. <http://cognitionandculture.net/webinars/speaking-our-minds-book-club/cats-tacs-and-kunvenshuns>.
- Scott-Phillips, Thom. 2015b. *Speaking Our Minds: Why human communication is different, and how language evolved to make it special*. London: Palgrave Macmillan.
- Selvanathan, Nagarajan. 2016. Specificational copular clauses as inverted predications with a semantics of equation. *University of Pennsylvania Working Papers in Linguistics*, **22**(1).
- Seuren, Pieter. 1998. *Western Linguistics: An historical introduction*. Oxford: Blackwell.
- Shanahan, Murray. 2016. The Frame Problem. In: Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, spring 2016 edn. Metaphysics Research Lab, Stanford University.
- Shannon, Claude E. 1948. A mathematical theory of communication. *Bell System Technical Journal*, **27**(3), 379–423.
- Shannon, Claude E., Weaver, Warren, & Burks, Arthur W. 1951. The mathematical theory of communication. *Philosophical Review*, **60**(3), 398–400.
- Sheehan, Michelle, & Hinzen, Wolfram. 2011. Moving towards the edge. *Linguistic Analysis*, **37**(3-4), 405–458.

- Smith, Barry. 2005. Against Fantology. *Pages 153–170 of:* Marek, Johann C., & Reicher, Maria E. (eds), *Experience and Analysis*. Vienna: HPT&ÖBV. Available at: http://ontology.buffalo.edu/bfo/Against_Fantology.pdf.
- Smith, Neil. 1999. *Chomsky: Ideas and Ideals*. Cambridge: Cambridge University Press.
- Smith, Robin. 2017. Aristotle's Logic. *In:* Zalta, Edward N. (ed), *The Stanford Encyclopedia of Philosophy*, spring 2017 edn. Metaphysics Research Lab, Stanford University.
- Sommers, Fred. 1969. Do We Need Identity? *Journal of Philosophy*, **66**(15), 499–504.
- Sommers, Fred. 2005. Intellectual autobiography. *Pages 1–25 of:* Oderberg, David S. (ed), *The old new logic: Essays on the philosophy of Fred Sommers*. Cambridge, MA: MIT Press.
- Sperber, Dan, & Wilson, Deirdre. 1995. *Relevance: Communication and cognition*. Oxford: Blackwell.
- Stalnaker, Robert. 1974. Pragmatic presuppositions. *Pages 197–213 of:* Munitz, Milton K., & Unger, Peter (eds), *Semantics and philosophy*. New York: New York University Press.
- Stowell, Tim. 1978. What was there before “there” was there. *Pages 458–471 of:* Farkas, Donka, Jacobsen, Wesley M., & Todrys, Karol W. (eds), *Papers from the fourteenth regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistics Society.
- Stowell, Tim. 1981. *Origins of phrase structure*. Ph.D. thesis, MIT.
- Stowell, Tim. 1983. Subjects across categories. *The Linguistic Review*, **2**, 285–312.
- Stowell, Tim. 1991. Small clause restructuring. *Pages 182–218 of:* Freiden, Robert (ed), *Principles and parameters in comparative grammar*. Cambridge, MA: MIT Press.
- Svenonius, Peter. 1994. *Dependent nexus: Subordinate predication structures in English and the Scandinavian Languages*. Ph.D. dissertation, UCSC, Santa Cruz, CA.
- Tarski, A. 1933/1983. The concept of truth in the languages of the deductive sciences. *Pages 152–78 of:* Corcoran, John (ed), *Logic, semantics, metamathematics, papers from 1923 to 1938*. Indianapolis: Hackett Publishing Company.
- Tattersall, Ian. 2002. *The monkey in the mirror*. Oxford: Oxford University Press.
- Tattersall, Ian. 2017. How can we detect when language emerged? *Psychonomic Bulletin & Review*, **24**(1), 64–7.

- Thornbury, Scott, & Slade, Diana. 2006. *Conversation: From description to pedagogy*. Cambridge: Cambridge University Press.
- Uckelman, Sara L. 2015. The logic of categorematic and syncategorematic infinity. *Synthese*, **192**(8), 2361–2377.
- Van Peteghem, Marleen. 1991. *Les phrases copulatives dans les langues romanes*. Wilhelmsfeld: gottfried egert verlag.
- von Heusinger, Klaus. 1999. *Intonation and Information Structure. The representation of focus in phonology and semantics*. Habilitationsschrift, Universität Konstanz.
- von Heusinger, Klaus. 2013. The salience theory of definiteness. In: Capone, A, Lo Piparo, F, & Carapezza, M (eds), *Perspectives on Pragmatics and philosophy*. Berlin: Springer.
- von Stechow, Arnim. 1996. The different readings of Wieder ‘Again’: A structural account. *Journal of Semantics*, **13**(2).
- Welch, Nicholas. forthcoming. Copulas and light verbs as spellouts of argument structure: Evidence from Dene. In: Arche, María J., Fábregas, Antonio, & Marín, Rafael (eds), *The grammar of copulas across languages*. Oxford: Oxford University Press.
- Wharton, Tim. 2003. Natural pragmatics and natural codes. *Mind & Language*, **18**(5), 447–77.
- Wharton, Tim. 2009. *Pragmatics and non-verbal communication*. Cambridge: Cambridge University Press.
- Whitehead, Alfred North, & Russell, Bertrand. 1963. *Principia Mathematica*. Cambridge: Cambridge University Press.
- Wiggins, David. 1965. Identity statements. In: Butler, R. J. (ed), *Analytical philosophy*. Oxford: Basil Blackwell.
- Wiggins, David. 1984. The Sense and Reference of predicates: A running repair to Frege’s doctrine and a plea for the copula. *The Philosophical Quarterly*, **34**(136), 311–28.
- William of Sherwood. 1968. *Treatise on Syncategorematic Words*. University of Minnesota Press. Norman Kretzmann, trans.
- William of Sherwood. 2012. *Syncategorematica*. Hamburg: Felix Meiner Verlag. C.Kann, R. Kirchhoff (eds), trans.

- Williams, C. J. F. 1989. *What is identity?* Oxford: Clarendon Press.
- Williams, Edwin. 1980. Predication. *Linguistic Inquiry*, **11**(1), 203–38.
- Williams, Edwin. 1983a. Against small clauses. *Linguistic Inquiry*, **14**(2).
- Williams, Edwin. 1983b. Semantic vs. syntactic categories. *Linguistics and Philosophy*, **6**, 423–446.
- Williams, Edwin. 1987. NP trace in theta theory. *Linguistics and Philosophy*, **10**(4), 433–447.
- Wilson, Deirdre, & Sperber, Dan. 2004. Relevance theory. *Pages 607–32 of: Horn, L. R., & Ward, G. (eds), The handbook of pragmatics*. Oxford: Blackwell.
- Wittgenstein, Ludwig. 1922. *Tractatus Logico-Philosophicus*. New York: Routledge and Kegan Paul. Trans: Pears, D.F. and McGuinness, B.F., 1974.
- Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Basil Blackwell.
- Woodard, Andrew. 2013a. *Moltmann's arguments that presentational pronouns are not ordinarily referring NPs*. MS, Durham University.
- Woodard, Andrew. 2013b. *Phase theory describes a working memory format*. MS, Durham University.
- Zalta, Edward N. 2011. Logic and Metaphysics. *In: Gupta, A., & van Benthem, J. (eds), Logic and Philosophy Today: Volume Two*. London: College Publications.
- Zalta, Edward N. 2017. Gottlob Frege. *In: Zalta, Edward N. (ed), The Stanford Encyclopedia of Philosophy*, spring 2017 edn. Metaphysics Research Lab, Stanford University.
- Zellou, Georgia. 2010. Tigrinya fronted copula constructions: Focus and evidence of speaker marking information relevant to hearer expectations. *Pages 106–21 of: Rice Working Papers in Linguistics*, vol. 2.