

# 5

## Pragmatism and Binding

Stephen Neale\*

### 1. Introduction

Names, descriptions, and demonstratives raise well-known logical, ontological, and epistemological problems. Perhaps less well known, amongst philosophers at least, are the ways in which some of these problems not only recur with pronouns but also cross-cut further problems exposed by the study in generative linguistics of morpho-syntactic constraints on interpretation. These problems will be my primary concern here, but I want to address them within a general picture of interpretation that is required if wires are not to be crossed. That picture will be sketched in sections 3 and 4; subsequent sections will focus on pronouns and binding, drawing heavily on what has preceded.

---

\* It will be obvious to linguists that this chapter is written by a philosopher, albeit one very sympathetic to the generative enterprise as articulated over the years by Noam Chomsky. Chomsky's ideas permeate both the syntactic and philosophical theses pushed here. As I was completing the chapter I was introduced to several works in linguistics I wish I had seen earlier, in particular Bresnan (2001), Elbourne (2001, 2002), Jacobson (1999, 2000), Kayne (2002), Keenan (2002), Reuland (2001*a,b*), Reuland and Sigurjónsdóttir (1997), Safir (forthcoming), and Szabolcsi (2003). I gesture towards some of these in various footnotes (Very likely I would have gone about certain matters a little differently if I had come across these works earlier; and I am not convinced some of the discussion could not be more usefully formulated within variable-free systems.) I thank the members of the NYU Linguistics Department for inviting me to present this work at their colloquium series—discussion (then and subsequently) with Mark Baltin, Paul Elbourne, Richard Kayne, Paul Postal, Anna Szabolcsi, and Michal Starke was most helpful. I also thank Laurien Berkeley, José Luis Bermúdez, Robert Fiengo, James Higginbotham, Paul Horwich, Mikael Karlsson, Richard Larson, Colin McGinn, Eric Reuland, Ken Safir, Stephen Schiffer, Barry Smith, Zoltan Szabó, Matthew Whelpton, Deirdre Wilson, and Höskuldur Þráinsson for their advice and suggestions, and Erlendur Jónsson, Hörður Sveinsson, Hrafn Ásgeirsson, Jón Ólafsson, Ólafur Páll Jónsson, Mikael Karlsson, Ósk Sturludóttir, Sigrún Svavarsdóttir, Ragna Guðmundsdóttir, and Þórdís Helgadóttir for help with some of the Icelandic examples. I gratefully acknowledge the generous support of Rutgers University, the University of Iceland, the Georg Brandes School, Institute for Nordic Philology, and the John Simon Guggenheim Foundation.

Interpreting an utterance or inscription involves substantially more than identifying and interpreting the individual words uttered and grasping their syntactic arrangement. Probably there isn't much it *doesn't* involve, and it is hardly surprising that we have not yet succeeded in producing a theory of interpretation with much empirical clout. There have been successes in some of the *subtheories*—phonology and syntax, for example. But there is widespread suspicion that producing an overarching theory of interpretation will require nothing short of a complete theory of mind.<sup>1</sup> Paul Grice once summed up the situation rather well to me. Compare the following 'maxims':

- (1) Put the subject first; put the object after the verb.
- (2) Be relevant; be informative.

We appear to have improved upon (1):

$$\begin{aligned} S &\rightarrow NP VP \\ VP &\rightarrow V NP. \end{aligned}$$

But, said Grice, we haven't got the foggiest how to improve upon (2).

It is not just in the realm of implied meaning (or implicature, to borrow Grice's term of art) that interpretation may seem wide open. Identifying who someone intends by 'John' or 'that man' or 'he', or 'his', for example, is not simply a matter of linguistic decoding; it involves the exercise of seemingly inferential abilities, and to this extent quite general concerns about interpretation impinge upon the identification of what a speaker is *saying* as well as what he is *implying*.

Where pronouns are concerned, traditional grammars suggest something like the following picture. Demonstrative pronouns ('this', 'that') are used to refer to things that are *salient* (or being made so by their users). Personal pronouns are used to refer to persons: first-person pronouns ('I', 'me', 'my', 'mine', 'myself') are used to refer to oneself; second-person pronouns ('you' etc.) are used to refer to one's audience; and third-person pronouns ('he', 'she', 'it' etc.) are used in place of 'fuller' phrases ('Plato', 'Athens', 'the current budget deficit') to refer to persons, places or things speakers could have referred to using those fuller phrases. The personal pronouns also have possessive ('his') and reflexive ('himself') forms; the third-person has an interrogative/relative ('who') form; and some differ in form according as they are subjects ('I', 'he', 'who') or objects ('me', 'him', 'whom'). Possessives indicate possession (construed broadly), reflexives signal anaphoric connections to expressions (of the same number and gender) in the immediate

---

<sup>1</sup> See, for example, Chomsky (2000), Davidson (1986), and Fodor (1983, 2001). For assessment, see Carston (2002).

environment, and the interrogative/relative form is used in asking questions ('who owns a donkey?') and in forming relative clauses ('every man who owns a donkey').

While there is much sense in this traditional taxonomy it is at best a first step in coming to terms with matters of interpretation familiar to philosophers and linguists who have wrestled with opacity, deixis, scope, variable-binding, and (my personal *bête noir*) co-reference. Although many philosophers share with linguists the desire to understand how pronouns fit into a general account of the workings of natural language, much of the interest in pronouns in philosophy is driven by the shenanigans these devices get up to in indexical inferences, modal statements, and propositional attitude ascriptions, and in talk of skepticism, self-knowledge, substitutivity, existence and identity, including personal identity—'Hesperus is Phosphorus, but John does not think *it* is', 'I thought your yacht was longer than *it* is'; 'every barber thought *his* reflection was of someone other than *him*'; 'Hob thinks a witch killed Nob's cow, and Bob thinks *she* blighted his mare'; 'if a man is born at the rising of the Dog-Star, *he* shall not die at sea'; 'most men who have only one donkey take good care of *it*'; 'every pilot who shot at *it* hit the MiG that was chasing *him*', '*I* will be buried tomorrow if, and only if, my body is buried tomorrow, yet *I* will be held in great esteem by my colleagues tomorrow even if my body is not.')'

Largely under Chomsky's tutelage, the accomplishments of empirical linguistics in the investigation of syntactic constraints on interpretation are impressive—despite many unsolved problems and despite the fact that every proposed theory of practically every phenomenon of significance appears to be frustrated by all sorts of counterexamples. If there is one overarching problem that has plagued the study of pronouns it is this: from the perspectives of syntax and interpretation, all sorts of superficial generalizations and distinctions can be made, but it is unclear which are of theoretical significance and which are mere distractions; and equally it is unclear which of a number of competing unifications and classifications will pay theoretical dividends and which are ultimately dead-ends. As a result, all sorts of problematic assumptions and misleading locutions are apt to burrow their way into the work of even the most philosophically attuned linguists, as well as into the work of the most linguistically attuned of those philosophers mad enough to get involved. Probably anyone bothering to read this paper knows, or has known, a philosophically attuned linguist or a linguistically attuned philosopher. But I fear interaction has not been good as it should have been, and the fact remains that getting a philosopher to give up a day's sailing on the Hudson to discuss long-distance reflexives in Icelandic, infinitival clauses,

subject-orientation and the Binding Theory is like trying to get a linguist to give up a day's sailing off Cape Cod to discuss privileged access, speaker's intentions, and epistemic asymmetry. And this is a crying shame, because on all of these issues, it seems to me, philosophers and linguists have an awful lot to gain from working together, making some of their work on these topics more accessible to the other party, and critiquing one another's works in helpful ways.<sup>2</sup>

It is in this spirit that the present chapter is written. It's hardly cutting-edge syntax or formal semantics, and it's hardly cutting-edge moral psychology or philosophy of mind. But it is a serious attempt to respond to pressures exerted by all four, rather than cutting-edge talking-down-to-philosophers-and-linguists-at-the-same-time. We are lacking some common solid ground, it seems to me, a place where linguists and philosophers can meet to address issues of mutual concern using words like 'pronoun', 'anaphor', 'binding', 'scope', 'syntactic', 'semantic', 'pragmatic', 'co-referential', 'interpretive', 'ambiguity', 'context', and 'identity' without driving one another to drink. Syntactic and interpretive notions are sometimes conflated in the literature, as are importantly distinct interpretive notions. And spells cast by false or insignificant generalizations only add to our woes. Some effort is required to tease apart things that have been run together and to recombine key elements in ways that make the issues sufficiently transparent and render the central theoretical posits truly explanatory, but I intend the whole of this chapter to be accessible to both linguists and philosophers, so members of one tribe may find themselves cruising through one section while members of the other are busy making extensive notes. This is unavoidable, I think, and I ask for readers' forbearance.

## 2. Pronouns as Variables

For many years, philosophical and logical interest in pronouns centred on distinguishing their use as something akin to the *bound variables* of quantification theory, their *deictic* (or *demonstrative*) use to refer to salient individuals ('*he* looks happy'), and their use as devices that obviated repetition ('if John does not hurry, he will miss the train').<sup>3</sup> The dictum 'to be is to be the value of a bound variable' probably does not hold as much sway as it did when Quine first pressed it, but the

---

<sup>2</sup> Matters are complicated by (i) the quite unwarranted hostility of many philosophers to Chomsky's talk of 'knowledge' of language, and (ii) the rightly sceptical position many linguists take towards some philosophers' talk of 'reference' which appears to have dragged in its wake an unwarranted scepticism about the notion of *speaker's* reference.

<sup>3</sup> Quine (1960), Geach (1962).

quantifier-variable combination is still part of philosophy's *lingua franca*. When we introduce variables in introductory logic classes we sometimes get the ball rolling by saying they are the formal counterparts of third-person pronouns. The following semester, we teach the philosophy of language, and if third-person pronouns come up in discussions of reference we begin with the statement that they are the natural language counterparts of variables in logic (raising a few eyebrows amongst those in the back row). The matter of the precise relationship between pronouns and variables is one among many that come up in the study of *anaphora* (from the Greek, "carry back"), construed as the study of *interpretive dependencies*, the study of the ways in which the interpretations of occurrences of certain expressions, particularly third-person pronouns, are tied to the interpretations of occurrences of other expressions. For several reasons, the relationship is nowhere near as straightforward as many philosophers supposed initially.

(i) Unlike variables in logic, natural language pronouns are typically marked for such things as *gender*, *number*, and *case*: 'she' is nominative, feminine, and singular; 'them' is accusative and plural. And it is a matter of some debate whether such features impinge upon the truth or falsity of what people say when they utter sentences containing them. Some pronouns are also singled out as *reflexive*, and it is matter of debate whether occurrences of 'he', 'him', 'his', and 'himself' impinge upon truth conditions in the same way.

(ii) Unlike variables in logic, natural language pronouns appear to bifurcate in respect of *locality* conditions on binding.<sup>4</sup> For example, 'him' *cannot* function as a variable bound by 'every man' in (1), whereas the so-called reflexive 'himself' *must* so function in (2):

- (1) every man loves him
- (2) every man loves himself.

That is, (1) *cannot*, but (2) *must*, be understood as expressing what (2') expresses:<sup>5</sup>

- (2') [every  $x$ : man  $x$ ]  $x$  loves  $x$ .

---

<sup>4</sup> Chomsky (1981, 1986, 1995).

<sup>5</sup> Throughout, the formalism of a system of restricted quantification RQ will be assumed in illuminating the 'logical forms' of sentences. Assuming a Tarskian account of quantification, the relevant axiom may take the form of (ii):

- (i) [every  $x_k$ :  $\phi(x_k)$ ]  $\psi(x_k)$  is true of a sequence  $s$  (of objects) iff  $\psi(x_k)$  is true of every sequence that  $\phi(x_k)$  is true of differing from  $s$  at most in the  $k$ -th position.

This interpretive contrast appears to involve *locality*: the non-reflexive ‘him’ is *too close* (in some sense to be elucidated) to ‘every man’ in (1), but not in (3), for example:

(3) every man loves the woman who married him.

And the reflexive ‘himself’ is *too far* from ‘every man’ in (4):

(4) every man loves the woman who married himself.

So, where (3) *can* be used to express what (4′) expresses, (4) *cannot*:

(4′) [every  $x$ : man  $x$ ]  $x$  loves the woman who married  $x$ .<sup>6</sup>

Since there is no condition on variables in first-order logic corresponding to this locality difference, its existence in natural language is something that needs to be described exactly and explained. For it *appears* that natural language is making things more complicated than it might.

(iii) Wherever we find a pronoun bound by a quantifier, we can replace the quantifier by a name (or some other singular term). For example, we get the following quantified-singular pair:

(5) every man loves himself

(5′) John loves himself

The naïve answer to the question of how the pronoun functions in the singular case (5′) is that it is co-referential with ‘John’, that it is a special type of referring expression whose reference is determined by some other expression. On such an account, pronouns may function in a way that variables in logic do not. If we are wedded to a choice between treating the pronoun in (5′) as a variable or treating it as a constant, it would seem we will have to go with the latter. And this might lead to the idea that ‘himself’ is a device for avoiding repetition, that an underlying form ‘John loves John’ surfaces as (5′) in ordinary English.<sup>7</sup>

But are we not missing something on such an analysis? Surely our capacity to understand sentences like (5′) is connected in some important way to our capacity

---

<sup>6</sup> If definite descriptions are quantifier expressions, we can get a little closer to the logical form of (4) with something like (i) (reflecting Geach’s (1972) *Latin prose theory* of relative clauses):

(i) [every  $x$ : man  $x$ ] [the  $y$ : woman  $y$  •  $y$  married  $x$ ]  $x$  loves  $y$ .

If, as Quine (1960) and Evans (1977) suggest, relative pronouns are devices of abstraction, we get still closer with something like (ii), where *who*  $z$  functions as an abstraction operator (see n. 9):

(ii) [every  $x$ : man  $x$ ] [the  $y$ : woman  $y$  •  $y$ (*who*  $z$ :  $z$  married  $x$ )]  $x$  loves  $y$ .

For immediate purposes, such a level of detail is not necessary.

<sup>7</sup> In the context of generative grammar, see Langacker (1966), Lees and Klima (1963), Postal (1966), and Ross (1967).

to understand sentences like (5).<sup>8</sup> Indeed, there is a strong intuition that ‘himself’ functions *identically* in (5) and (5’), that a single closed predicate ‘loves himself’— $\lambda x(x \text{ loves } x)$ —occurs in both sentences, that (5) and (5’) are both used to make claims involving the following condition:<sup>9</sup>

(5'')  $x \text{ loves } x$ .

(5) is used to say that this condition is true of every man, and (5’) is used to say that it is true of John. Isn’t it this that explains the validity and immediacy of the following inference?

(6) every man loves himself; John is a man; so John loves himself.

So if ‘himself’ is really functioning as a bound variable in the quantified case (5)—no other theory seems to be lurking in the wings—shouldn’t we at least countenance the idea that it also functions in this way, rather than as a device of co-reference, in the singular case (5’)? And if that is right, in order to preserve as much strict compositionality as possible, shouldn’t we see all predicates as ultimately devices of abstraction, ‘snores’ expressing  $\lambda x(x \text{ snores})$ , ‘loves’ expressing  $\lambda y(\lambda x(x \text{ loves } y))$ , and so on? We cannot be sure yet that there is not a genuine *ambiguity* between a bound variable reading of ‘himself’ and a co-referential reading, one that is truth-conditionally inert in (5’) but not in more interesting examples such as those involving verbs of propositional attitude. But we have to start somewhere, let us assume until further notice that we have only one anaphoric reading of (5’) to deal with: ‘himself’ is bound by ‘John’ in (5’) in whatever way it is bound by ‘every man’ in (5), details to be provided.<sup>10</sup>

(iv) It is evident that that variable-binding cannot supply everything we need. Take (7) and (7’):

(7) every man loves his wife

(7’) John loves his wife

<sup>8</sup> Evans (1980), Higginbotham (1980, 1983a).

<sup>9</sup> Geach (1962, 1972), Heim (1982/1988), Heim and Kratzer (1998), Kamp (1981), Partee (1975), Reinhart (1983), Salmon (1986, 1992), Soames (1990, 1994), Wiggins (1976).  $\lambda$  is the *lambda* (or *abstraction*) operator. On the usage adopted here,  $\lambda x(x \text{ snores})$  and  $\lambda x(x \text{ loves } x)$  are one-place predicates. Thus  $John(\lambda x(x \text{ loves } x))$  is a sentence.

<sup>10</sup> I say I am going to start here, but I am also going to *finish* more or less here. The combined force of points made by Castañeda (1966, 1967, 1968), Geach (1962), Heim and Kratzer (1998), Partee (1975), Reinhart and Reuland (1993), Reuland (2001a,b), Salmon (1986, 1992), and Soames (1990, 1994) seems to me to undermine beyond the point of possible repair any unitary *de jure* co-reference account of anaphora such as the one offered by Evans (1977, 1980), and any ‘pragmatic’ account of the phenomena, such as those proposed by Lasnik (1976) and Bach (1987, 1994). Some of these points will emerge as we proceed.

On one use of ‘his’ it is a device of anaphora, bound by ‘every man’ in (7), and by ‘John’ in (7’), the respective sentences used to make claims involving the following condition:

(7'')  $x$  loves  $x$ 's wife.<sup>11</sup>

(7) is used to say the condition is true of every man; and (7') to say it is true of John, the common predication being  $\lambda x(x \text{ loves } x\text{'s wife})$ . In line with what was said about reflexives, let us call this the *bound* use of ‘his’.

On a second use of ‘his’, it is *free*: it is used to make independent, indexical reference to some particular male. Suppose we have been talking about Paul, who is known to be married to a woman who is utterly captivating. We might use ‘his’ in (7) and (7') to refer to Paul, to make claims involving not condition (7'') but condition (7'''):

(7''')  $x$  loves Paul's wife.

If we insist on stating everything in terms of variables, then we can say that on this use of ‘his’ it functions as a *free* variable, an expression to which some specific value must be assigned for interpretation to take place.

A question now arises concerning the precise relationship between these two uses of ‘his’. One might rest content with an ambiguity story and cite in its defence the fact that translating (7) and (7') into some languages requires making a choice between two quite distinct pronouns depending upon whether the bound or indexical reading of the English ‘his’ is intended.<sup>12</sup> Alternatively, one might

---

<sup>11</sup> Again, I suppress some logical structure. If possessives like ‘his wife’, ‘John’s wife’, and ‘that man’s wife’ are definite descriptions, the condition given in (7'') might be specified more fully using (i):

(i) [*the y: y wife x*]  $x$  loves  $y$ .

I will offer something more appealing later.

<sup>12</sup> In Icelandic, for example, a possessive pronoun agreeing in gender (feminine) and case (accusative) with the noun for ‘wife’ is used for the bound reading:

(i) [*sérhver maður*]/*Jón elskar konuna sína*  
[every man]/John loves wife-DEF POSS-FEM-ACC  
[every man]/John loves his (bound) wife.

(The letter *ð* (*eð*, ‘eth’) is pronounced like ‘th’ in English ‘the’.) But if the English ‘his’ is understood as an indexical it must be translated using a simple genitive *hans* here, which occurs in the masculine and enters into no agreement relations whatsoever with *konuna* (‘wife’):

(ii) [*sérhver maður*]/*Jón elskar konuna hans*  
[every man]/John loves wife-DEF his-MASC-GEN  
[every man]/John loves his (free) wife

The question would now arise whether there are languages that distinguish bound and indexical uses of ‘he’ and ‘him’, which Icelandic does not. I should stress that matters are more complex in Icelandic than this note might suggest. In order to add a modicum of breadth to the discussion and motivate certain avenues of research, later in the chapter I shall make frequent comparisons between English and Icelandic drawing upon published work by generative linguists (see below for



eschew the aforementioned ambiguity theory and seek some form of *unification*, taking the model of bound and free occurrences of variables at face value: a variable must be assigned a value somehow, and this means assigning it a single value directly or allowing its value to vary systematically with the value of some other expression.

(v) Closely related to the matter of ambiguity is the matter of the interpretation of expressions pared down by rule-governed ellipsis.<sup>13</sup> We need to explain the contrast between (9) and (10), for example:

(9) John loves himself. So does Paul.

(10) John loves his wife. So does Paul.

There is a single reading of (9). But there are two quite distinct readings of (10), one upon which Paul is being said to satisfy (10'), another on which he is being said to satisfy (10''):

(10') *x loves x's wife*

(10'') *x loves John's wife*

This might suggest we need to distinguish a reading of (10) upon which 'his' is bound by 'John' and another upon which it is an indexical used to refer to John, the same person 'John' is being used to refer to. This would comport with the idea that the ellipted structure is understood as 'so does Paul love his wife' with 'his' preserving its interpretation either as a variable to be bound by the subject expression or as an indexical used to refer to John. It would also explain why (9) has only one reading: reflexives permit only bound readings.<sup>14</sup>

But might the second reading of (10) not be the product of a *referentially dependent* use of 'his'? The idea would be that on this second reading 'his' is not bound by 'John' in the first clause of (10'') but is rather co-referential with 'John' by virtue of being referentially dependent upon it. On such an account, the ellipted structure is understood as 'so does Paul love his wife' with 'his' preserving its interpretation as referring to whoever 'John' refers to.

---

references) and investigations I have undertaken myself over the past few years with the help of various Icelandic philosophers and friends.

<sup>13</sup> Heim and Kratzer (1998), May (2002), Partee (1975), Sag (1976), Williams (1977).

<sup>14</sup> This proposal and the common predication in singular and quantified pairs seems to explain the fact that we understand and immediately see the validity of things like (i):

(i) Every man loves himself. Since John is a man, it follows that John loves himself.  
And since Paul is a man, it follows that he does too.

(vi) A pronoun may be used in such a way that it seems to be interpreted neither as a variable bound by some antecedent phrase nor as an indexical.<sup>15</sup> That is, anaphoric dependencies seem to exist that do not involve binding. Consider (11) and (11'):

- (11) Just one man ate haggis. He was ill afterwards.  
 (11') John ate haggis. He was ill afterwards

Treating 'he' as a variable bound by 'just one man' in (11) yields the wrong result:

- (12) [*just one x: man x*] (*x ate haggis • x was ill afterwards*).

(12) can be true if two men ate haggis but only one of them was ill afterwards, i.e. its truth is consistent with the *falsity* of the first conjunct of (11).<sup>16</sup> Thus (12) appears to capture the meaning not of (11), but of (12'):

- (12') Just one man ate haggis and was ill afterwards.

So we are still without an account of the pronoun in (11). However, the fact that (11) and the ungainly (11'') below appear to be equivalent suggests there is something to be gained by exploring the view that the pronoun in the former functions as a disguised definite description:<sup>17</sup>

- (11'') Just one man ate haggis. The man who ate haggis was ill afterwards.

But the desire to preserve a uniform treatment of pronouns as variables might suggest exploring a rather different idea: that the occurrence of 'he' in (11) is a variable bound by *something other than* 'just one man', something not revealed until we have an account of the underlying *logical form* of the sentence and a general theory of the *pragmatics of discourse*.<sup>18</sup>

Now what of the singular case (11')? If the quantifier 'just one man' cannot bind the pronoun 'he' in (11), there is no good reason to think 'John' can in (11'). We would appear to have three options for dealing with the pronoun in (11') we could view it (i) as an indexical being used to refer to John, this being licensed in some way by John's salience; (ii) as a pronoun of laziness, interpreted as if it were just another occurrence of 'John'; (iii) as a device of co-reference—the difference between positions (ii) and (iii) would need to be articulated clearly.

The issue raised by the pronoun in (11) recurs with more interesting examples. Consider the so-called 'donkey' sentence (13):

<sup>15</sup> Cooper (1979), Evans (1977), Heim (1982/1988), Kamp (1981), Partee (1972).

<sup>16</sup> Evans (1977).

<sup>17</sup> Cooper (1979), Davies (1981), Elbourne (2001, 2002), Evans (1977, 1980), Neale (1990, 1994).

<sup>18</sup> Heim (1982/1988), Kamp (1981).

(13) every man who bought just one donkey paid cash for it.

If the quantifier ‘just one donkey’ is to bind ‘it’ it will have to be given large scope:

(14) [*just one y: donkey y*] [*every x: man x • x bought y*] *x paid cash for y*.

But again this yields the wrong result. The fact that (13) appears to be equivalent to the ungainly (13') below suggests there is something to be gained by exploring the view that the pronoun in (13) functions as a disguised definite description:<sup>19</sup>

(13') every man who bought just one donkey paid cash for the donkey he bought.

On such an account, the description the donkey pronoun ‘it’ abbreviates itself contains a pronoun ‘he’ bound by the subject expression ‘every man who bought a donkey’. Whilst this is certainly unproblematic as far as (13') is concerned, it raises questions about the nature of the mechanisms involved in interpreting a pronoun that appears to be anaphoric but not bound by the expression upon which it appears to be anaphoric, and this might suggest sweeping up all anaphoric pronouns within a theory of the pragmatics of discourse.<sup>20</sup> Alternatively, it might suggest sweeping them up with a description-based approach.<sup>21</sup>

With these logical issues in mind, let us now turn to the general interpretive framework within which they should be soluble.

### 3. Linguistic Pragmatism

Let us call a theory that aims to explain how hearers manage to identify what speakers are seeking to communicate a *theory of utterance interpretation*, or a *theory of interpretation* for short. To say that we are interested in providing a theory of interpretation is not to say we are prejudicing the issue against communication that does not involve speech or writing. It might simply turn out—it surely will—that a theory of interpretation will make reference to cognitive capacities involved in interpreting non-linguistic acts of communication, indeed non-linguistic acts more generally.<sup>22</sup>

---

<sup>19</sup> See n. 17.

<sup>20</sup> Heim (1982/1988), Kamp (1981/1984).

<sup>21</sup> Elbourne (2001, 2002).

<sup>22</sup> See Grice (1989), Sperber and Wilson (1986, 1995), Carston (2002).

The project of explaining interpretation has many components and involves people from several fields. Philosophers have two rôles, one in the board room, the other on the shop-floor, as it were. First, they will attempt to articulate clearly the nature of the project (distinguishing it carefully from various other projects with which it might be confused), distinguish clearly the various sub-projects, and distinguish and analyse the central concepts or at least the relations between them (for example, meaning, saying, implying, referring, and intending). At the same time, they will attempt to work alongside linguists whose expertise involves explaining how individual words are assembled into sentences and the extent to which communicatively relevant features of the sentences we use to say things depend upon features of the words out of which they are assembled and the mode of assembly itself. And alongside psychologists who can tell philosophers and linguists about cognition, in particular about the way we integrate information from different sources and channels in the process of identifying what someone is trying to communicate.

Before getting down to pronouns *per se*, I want to sketch the general picture within which I think we should be operating. With distinct nods to the American Pragmatists, to Wittgenstein, Sellars, and Quine, to Perry, and to Sperber and Wilson, I call the general outlook I have on the matter of interpretation, *linguistic pragmatism* (or *pragmatism* for short). It is an outlook that can be held by philosophers, linguists, psychologists, and no doubt others—no diplomas are checked at the door—who take themselves to be involved in the project of constructing a general theory of (utterance) interpretation, construed as an empirical theory and, as such, a contribution to cognitive psychology. It might be seen as a collection of theses that can emerge *only* in the context of attempting to articulate the outlines of such a theory, theses whose truth may well have repercussions elsewhere but which are not themselves motivated by the desire to bolster this or that philosophical or political doctrine, even if distinctions inherent in certain theses emerged from reflections on the language used in stating philosophical positions and problems. (The pragmatist outlook may well be implicated in various works of a ‘contextualist’ nature, but I am anxious to distance my own views from extant contextualist proposals in epistemology, metaphysics, ethics, and political philosophy, many of which seem to me rather suspect).

Some of the central tenets of linguistic pragmatism were accepted by a number of British philosophers in the 1950s, particularly Austin, Strawson, and (contrary

to the claims of some pragmatists) Grice.<sup>23</sup> But it was not until the late 1970s, by which time the Language of Thought hypothesis articulated by Fodor, the Chomskyan idea of LF as a level of linguistic representation, and important distinctions made by Grice and Searle had truly sunk in, that the conceptual resources were generally available to articulate the outlook clearly and in a form that made it relevant to more formal studies of language that were by that time blossoming in linguistics and philosophy departments in the United States. To the best of my knowledge, it was not until the work of Sperber and Wilson began to appear in print in the early 1980s that pragmatists made sustained efforts to render explicit the basic tenets of their work. Indeed, without Sperber and Wilson's work, and the work of Chomsky, Fodor, Grice, and Searle upon which it drew, philosophy and linguistics might still lack the distinctions and resources needed to say anything more substantive than the ramblings about 'contextual meanings' and 'relative meanings' that issue periodically from the darker areas of philosophy and linguistics departments (not to mention departments or 'programs' housing people unaccountably known as 'theorists' or 'philosophers').

It would be a mistake, I think, to attempt a *definition* of linguistic pragmatism as it is essentially an outlook that engenders a very practical approach to interpretation. I cannot go into the sort of detail I go into in *Linguistic Pragmatism* here, so I have produced twenty-four labelled points (without much in the way argument) to give the general flavour of pragmatism as I think it should be developed. Some of the points are quite general or intuitive, others are very specific or theory-laden. Some are less central than others and could be withdrawn without upsetting the whole too much, but I am strongly inclined to go along with the whole lot, and there is no doubt that many gain strength through association with others. A few are held by some philosophers and linguists I would call anti-

---

<sup>23</sup> A somewhat simplistic picture of the relationship between the focus on 'ordinary' language and the use of 'ideal' or 'formal' languages appears to be accepted by many linguists and even some philosophers. The received view in linguistics appears to be that for some years there was a major philosophical conflict (between 'formalists' and 'informalists') which Grice somehow dissipated by distinguishing what a speaker said from what he 'conversationally implicated'. (The picture is perhaps fostered by a naïve reading of the opening paragraphs of Grice's 'Logic and Conversation' and by Strawson's (1969) bizarre claims about a 'Homeric struggle' in his inaugural lecture 'Meaning and Truth'.) Some of the people in the grip of this picture have been led to conclude that Grice was not actually a pragmatist. I know from conversations with him (a) that he saw the problem of providing an accurate account of what the speaker says when using an incomplete description as providing powerful *evidence* for pragmatism, and (b) that he never intended to be seen as denying pragmatism.

pragmatists; several are rejected by some people with pragmatist outlooks; and a few are conspicuous here by their absence elsewhere in the literature.<sup>24</sup>

1. *Cooperation*. Typically speakers (writers) want to be understood, and hearers (readers) seek to understand. To this extent they are involved in a cooperative exercise. *Ceteris paribus*, both parties tacitly assume they are using words with shared meanings, combining these words in accordance with a shared syntax, and operating in accordance with shared and very general, rational principles of *interpretation*, and enough shared beliefs to make this all worthwhile.

2. *Meaning*. A theory of interpretation should explain how hearers (and readers) manage to integrate linguistic and non-linguistic information to identify what a speaker (or writer) *meant* on a given occasion by uttering (or inscribing) a linguistic expression *X*.<sup>25</sup> Valuable information can be gleaned from examining situations in which we report on speech acts using sentences of the form,

by uttering (or writing) *X*, *A* meant that *p*

where the reporter is using the expression replacing '*A*' to pick out an agent and the expression replacing '*X*' to pick out a linguistic expression, and where the expression replacing '*p*' is a declarative sentence. Examples:

---

<sup>24</sup> Putting aside differences of terminology and philosophical temperament, as well as apparent disagreements about particular analyses, I am inclined to view all of the following as operating in a broadly pragmatist spirit—although my own brand of linguistic pragmatism is, I suspect, rather too ascetic, inferential, beholden to ordinary language strictures, and driven by underlying concerns about practicality and the concepts of society and regulation for many of them: Austin (1962), Bezuidenhout (1997), S. Blackburn (1984), W. Blackburn (1987), Blakemore (1987, 2002), Barwise and Perry (1983), Carston (1988, 1993, 2002), Chomsky (1976a, 1986, 1995, 2000, 2002), Crimmins (1992), Crimmins and Perry (1989), Evans (1982, 1985), Heal (1997) Fodor (1987, 2001), Grayling (1995), Grice (1989), Neale (1990, 1993), Quine (1940, 1960), Papafragou (1998*a,b*), Perry (1986, 1998, 2000, 2001), Récanati (1987, 1989, 1993, 2001), Rouchota (1992), Searle (1969, 1979, 1980), Sellars (1954), Sperber and Wilson (1986, 1995), and Strawson (1950, 1952). I suspect pragmatism is also taken for granted by many others who take truth to be property of what a speaker says or expresses by uttering a sentence *X* on a specific occasion, and not of what *X* itself says or expresses relative to a context. I shall point to or draw liberally from the work of these pragmatists as I go. By saying I am a pragmatist, I do not mean to be saying that I endorse the details of Récanati's (1989, 1993) or Rouchota's (1992) or Sperber and Wilson's (1986, 1995) or Bezuidenhout's (1997) or Carston's (2002) pragmatist analyses of definite descriptions and the attributive-referential distinction, all of which I find problematic. I propose an account that should be attractive not only to them but to Russellians and ambiguity theorists alike in Neale (2004, forthcoming a).

<sup>25</sup> Linguistic pragmatism does not necessarily assume there is much chance of ever producing an empirically interesting theory of interpretation. At least two pragmatists, Chomsky (2000) and Fodor (1983, 1987, 2001), have argued that asking for a theory of interpretation is tantamount to asking for a 'theory of everything', a complete cognitive psychology, because virtually anything can impinge upon the holistic process of interpretation. For more optimistic pragmatist outlooks, see Sperber and Wilson (1995, 1996) and Carston (2002). The present essay assumes neither outlook, but it brings into sharp relief the need for a clear picture of what the more tractable sub-theories of a theory of interpretation are supposed to do and how they must come together.

by uttering, 'I'm tired', John meant that he was tired

by uttering, 'I'm tired,' John meant that we were to leave.<sup>26</sup>

We should be suspicious of locutions of the form '*X* means that *p*', where the expression replacing '*X*' is being used to pick out a sentence (e.g. 'the sentence 'snow is white' means that snow is white').

3. *Explanation*. To interpret is to provide an *explanation*, and the concept of interpretation makes no sense in the absence of a *problem* to be solved. We reflexively generate *hypotheses* about the things we perceive. Nowhere is this more in evidence than when we perceive one another's *actions*. We act out of *reasons*. To interpret an action is to form a hypothesis about the *intentions* behind it, the intentions that *explain* it. Interpreting a speech act is a special case of this. The use of language is one form of rational activity, and the principles at work in the interpretation of linguistic behaviour are intimately related to those at work in interpreting intentional *non-linguistic* behaviour. What makes interpreting a speech act special is that a proprietary body of information, knowledge of language, is accessed immediately in the interpretation process. The hearer's or reader's goal is to identify what the speaker or writer meant. When this has been done, the interpretive problem has been solved.<sup>27</sup>

4. *Asymmetry*. The epistemic situations of the speaker and hearer are fundamentally asymmetric: the speaker *knows* what he means whereas the hearer has to *work it out*. If you want to find out whether I'm hungry (or in pain) you will have to watch me, see what I do, or ask me. *I* don't have to do that. I have 'privileged access' to that information.<sup>28</sup> Similarly if you want to know whether I am worried about missing my flight, where on an aeroplane I prefer to sit, or whether I think Norway is a member of the European Union. And similarly where we have speech. Unlike you, I have privileged access to what I mean when I utter *X* on a given occasion. We can characterize a typical speech situation as follows.

---

<sup>26</sup> The analysis of statements of the is general form was inaugurated by Grice (1957). See also Grice (1989), Strawson (1971), Schiffer (1972), and Neale (1992). For worries, see Blackburn (1984) and McDowell (1985).

<sup>27</sup> I mean this to apply equally to interpretations in literary theory. The idea of textual interpretation makes no sense if there is no problem (about, for example, a word or phrase, a character, a plot, a work, or even a whole genre) to which an interpretation constitutes a possible solution. It is embarrassing that some 'theorists' who also call themselves 'philosophers' are unable to see this.

<sup>28</sup> There are philosophers concerned to deny this idea today, but no coherent case has been made (indeed could be made) for a total failure of asymmetry. The idea of 'privileged access' to a state is often introduced with the idea of 'in corrigibility' (the idea that I cannot be mistaken about whether I am in the state in question). Whilst the case for denying this might be more promising, the fact that incorrigibility is at least *arguable* (and has been argued) in the first-person case, and is not in the least arguable (and has not been argued except, perhaps, by the deluded) in the third-person case, is enough to distinguish the cases here.

Person *A* intends to communicate something to some other person *B*. He selects a form of words *X* that he thinks will, in the circumstances, get across his point (and, perhaps, also get it across in some particular way or other. *A* knows what he means by uttering, ‘That’s his bank,’ for example. He knows *which thing* he meant by ‘that’, *who* and *what relation* he meant by ‘his’ and *what* he meant by ‘bank’.

*B*’s situation is quite different: *B* is trying to work out what *A* meant and he must use anything he can get his hands on to get the job done since he has no direct access to *A*’s communicative intentions. The words *A* uses constitute partial evidence for what *A* meant. Other evidence may come from the physical environment, from *B*’s take on the conversation up to that point (if any), from *B*’s beliefs about *A*, and a whole lot more besides. The epistemic asymmetry of speaker and hearer underscores (i) the need to separate the *metaphysical* question concerning what *determines* (or *fixes*) what *A* means and the *epistemological* question concerning what is used by others to *identify* what *A* means, and (ii) the need to scrutinize simplistic appeals to *contexts*, *maxims of conversation*, *salience*, and *pragmatic factors*, which are frequently (and mistakenly) introduced together with *intentions* in contemporary discussions as if these things conspire to bridge certain interpretive gaps. Scanning the context of utterance for salient objects and bringing to bear pragmatic principles (e.g. Grice’s conversational maxims) is not going to provide *A* with any information that will help him identify what he meant. From *A*’s perspective, context and pragmatic principles have already fulfilled their rôles: *A*’s perception of the context—whatever a context turns out to be—his perception of *B*’s perception of the context, the assumption that *B* is operating in accordance with the same pragmatic principles as *A*, and *A*’s estimation of *B*’s ability to work things out (and probably a whole lot more besides) have already impinged upon whatever processes led *A* to use the particular form of words he used with the particular intentions with which he used them.

5. *Reciprocity*. Despite the epistemic asymmetry, the perspectives of *A* and *B* are not independent. The asymmetry is *reciprocal* or *complementary* as in adjoining pieces of a jig-saw puzzle. In producing his utterance, *A* relies on what he takes to be *B*’s capacity to identify what he intends to convey; *B* assumes that *A* is so relying. And, possibly, so on. The ways in which *A* and *B* operate form a *dovetail joint* and are *mutually sustaining*. And to this extent, there is simply no possibility of making sense of *B*’s capacity to interpret *A* without making sense of *A*’s capacity to exploit that capacity, and vice versa. So the project of constructing a theory of interpretation may be approached from either of two complementary perspectives, and an adequate answer must make sense of both.



6. *Intention.* There is no plausible alternative to construing what *S meant* by uttering *X* on a particular occasion—assuming this to be the notion at the heart of a theory of interpretation—as determined by, and only by, certain very specific *interpreter-directed intentions* *A* had in uttering *X*. The precise content of a psychological state such as a belief or intention may be determined, in part, by something external to *A* and beyond *A*'s control ('externalism'). Furthermore, the formation of genuine intentions is severely constrained by beliefs. I cannot intend to become a prime number, intend to digest my food through my lungs on alternate Tuesdays, or swim from New York to Sydney because (roughly) I cannot intend *what I believe to be impossible*. (There is no need to get into the exact force of the modal or the exact formulation of the constraint here. It is enough to recognize, as Grice (1971) does, that it is severe.) If, as Grice suggests, what *A* meant by uttering *X* on a given occasion is determined by certain interpreter-directed intentions, then assuming he is being co-operative *A* cannot mean that *p* by uttering some sentence *X* if he believes it is impossible for his audience *B* (or at least any rational, reasonably well-informed interpreter in *B*'s shoes) to construe him as meaning that *p*. Among the things constraining *A*'s communicative intentions are *A*'s beliefs about the world, his (tacit) beliefs about the sorts of interpretive principles *B* will be employing, and his (tacit) estimation of *B*'s capacity to work certain things out (the list is not meant to be anywhere near exhaustive). So without some stage-setting *A* cannot mean that Jones is no good at philosophy by producing the sentence 'Jones has excellent handwriting and is always punctual', for example, or by reproducing the mating call of some exotic bird.

7. *Factorization.* What *A* meant by uttering *X* may be factored into what *A* said (or asked) by uttering *X* and what *A* only implied.<sup>29</sup> Thus, again following Grice, what *A* said and what *A* implied are determined by, and only by, certain very specific interpreter-directed intentions *A* had in uttering *X*.<sup>30</sup> Nonetheless, although

<sup>29</sup> In my view it is vital to distinguish between (roughly) (i) what *A* implied by uttering *X*, and (ii) what *A* implied by saying what he said. This distinction seems to me to get to the heart of Grice's distinction between conventional and conversational implicature, and to provide the framework within which to solve problems concerning the former (cf. Frege's notion of tone or colouring), non-detachability, dictiveness and formality, and central vs. non-central speech acts.

<sup>30</sup> Saul (2002a, 2002b) denies the existence of an exhaustive factorization of this type in Grice's work. The "common understanding of Grice," says Saul, is "unsustainable." (2002a: 228). "The problem for [it] lies in the fact Grice's characterizations of speaker meaning and conversational implicature are cast in very different terms—the former completely in terms of speaker intentions and the latter incorporating a good deal about the audience. As a result, the notions do not fit neatly into the simple picture I was taught" (2002a 228-9). "For Grice, what speakers say and what speakers implicate is not simply a matter of what they intend." (2002a: 229). "Conversational implicatures . . . are entirely removed from the control (or possibly even awareness) of speakers." (2002a: 230). Saul bases these and other claims on rather blinkered readings of selected passages in Grice's published work and appears to overlook (in both the readings and the consequent arguments) the importance of

it would be perverse to insist upon a distinction between what *A* meant and what *A* intended to mean (and for good reason if Grice is right), a distinction between what *A* said and what *A* intended to say is not one obviously lacking a point. In the first instance, we should separate (i) *what A intended to say by uttering X on a given occasion*, and (ii) *what a rational, reasonably well-informed interpreter in B's shoes would think A intended to say by uttering X on that occasion* (which is not to say there are not problems with the idea of a *rational, reasonably well-informed interpreter in B's shoes*). In cases where (i) = (ii), we can talk freely about *what the speaker said*. In cases where (i) ≠ (ii), certainly we *could* argue about which of (i) or (ii) or some third thing has the 'right' to be called *what is said*, but what would be the point? First, what third thing distinct from (i) and (ii) could be of any significance to a theory of interpretation? There is simply no rôle for a transcendent notion of what is said upon which (i) and (ii) converge when all goes well. It is the *coincidence itself* of (i) and (ii) that constitutes success, and it is the *potential* for such coincidence, independently of some third thing, that *gives sense* to the very idea of saying. (Contrary to linguistic appearances the concept of *intending* to say is, in fact, more basic than the seemingly simpler concept of saying.) Second, why is a choice between (i) and (ii) needed in cases where (i) ≠ (ii)? Conceptually they are distinct, and *both* are needed in a theory of interpretation. When all goes well, they coincide, and it's just too bad they don't always do so. There is no philosophical payoff in bestowing the honorific 'what was said' on one rather than the other, or on some third thing, when they diverge, (Everything I just said about saying carries over *mutatis mutandis* to implying and referring.)

---

the following to Grice's project: (i) the *cooperative nature* of typical talk exchanges and its rational underpinnings; (ii) the *epistemic asymmetry* of speaker and hearer, which engenders different theoretical locutions when switching perspectives; (iii) the *reciprocal* or *dovetailed nature* of these perspectives, including the fact there is meant to be a default assumption that, in the absence of evidence to the contrary, (a) the audience presumes the speaker to be operating in accordance with the Cooperative Principle, and (b) the speaker presumes the audience to be so-presuming; (iv) the *severe constraints beliefs place* on the formation of intentions; and (v) the fact that the communicative intentions Grice is concerned with (so-called M-intentions) are (a) *audience-directed*, (b) intended to be recognised, and so (c) severely constrained by the speaker's beliefs about the *audience* towards whom the intentions themselves are directed, in particular his beliefs about the likelihood of his intentions *being recognized* by the audience. (What the audience is to think and infer is an important part of the full story about the speaker's communicative intentions: what the speaker can M-intend is largely determined by what he can expect his audience to think and infer. (The last remark was added at Stephen Schiffer's suggestion. The wording is his.)) For discussion, see *Linguistic Pragmatism*.

On a related note, if what *A* said is determined by, and only by, certain very specific *interpreter-directed intentions* *A* had in uttering *X*, then at least some contemporary talk of 'contexts' 'fixing' or 'determining' aspects of what *A* said—for example the references of indexical expressions—must involve some form of confusion. See below.

Among the things that constrain the formation of *A*'s *saying-intentions* are *A*'s beliefs (including beliefs about his audience), his knowledge of the meanings of the words he is using, and his (tacit) knowledge of the syntax of the language he is using. Thus *A* cannot (intend to) say that snow is white by uttering the sentence 'grass is green', for example. And he cannot (intend to) say that John asked his brother to shave him by uttering 'John asked his brother to shave himself'. More generally, he cannot (intend to) say that *p* by uttering *X* if he believes it is impossible for his audience *B* (or at least any rational, reasonably well-informed interpreter in *B*'s shoes) to construe him as intending to say that *p*.

8. *Speakers*. Saying and implying are things *people* do. Following ordinary usage, *the speaker* is taken to be the understood subject, so to speak, of the verbs 'say' and 'imply' the verbs in talk about 'what is said' and 'what is implied.' (Similarly, with verbs such as 'communicate', 'convey', and 'get across'.) We should be initially suspicious of talk about what *uses of sentences* say (imply, communicate, etc.) and talk about what *sentences-relative-to-contexts* say (imply, communicate, etc.), unless such talk is taken to be straightforwardly translatable into talk about things that speakers are doing. (It is always a dead giveaway that a writer has precious little understanding of Grice's work when one finds the writer talking about what a *sentence* conversationally or conventionally implicates, a way of talking that is utterly alien to Grice and his project.) And we should deplore the unannounced slipping and sliding back and forth between different subjects of 'say' ('imply', 'communicate', etc.). At the same time, we should be open to the idea that new, *technical* uses of the verbs 'say' ('imply', 'communicate', etc.) may need to be defined, or at least developed, in the course of our inquiries, such uses earning their keep because of ineliminable theoretical work they do.

9. *Truth*. What *A* says and implies are the sorts of things that are true or false. (Perhaps *A* may say things that are neither true nor false, but it might prove useful to start out sceptical about this.) It does not follow that when *A* utters a sentence he says only one thing or that he implies only one thing. Nor does this talk of truth mean that in order to produce a theory of interpretation we shall have to construct a theory that recursively assigns *truth conditions* to sentences relativized to contexts of utterance (a *semantic* theory, in *one* sense of 'semantic'), or construct a theory that assigns things in the world to linguistic expressions relativized to contexts of utterance (a theory of *reference*, in *one* sense of 'reference').

10. *Judgment*. Our intuitive judgments about what *A* meant, said, and implied, and judgments about whether what *A* said was true or false in specified situations constitute the primary data for a theory of interpretation, the data it is the theory's

business to explain. (Since *no-one* has intuitive judgments about *what is said by a sentence X* relative to a context *C* or about the *semantic content* of *X* relative to *C* (these being philosophers' notions), several distinct mistakes would be involved in the claim that linguistic pragmatism aims to show that our intuitive judgments about what a speaker said may be 'unreliable guides to semantic content.' If talk of the 'semantic content' of a sentence *X* relative to a context *C* is just a snazzy way of talking about what the speaker said by uttering *X* on a particular occasion—the occasion that *C* is being used to partially model—then of course we can accept its empirical significance. If it is not, then its empirical significance must be justified in some other way, from *within* the theory of interpretation by reference to some empirical rôle it is required to play in an explanation of what a speaker says and implies by uttering *X* on a given occasion, in much the same way that notions such as scope and binding are motivated from within.

11. *Reference*. Saying typically involves *referring* and *saying of*. That is, saying something typically involves *referring to* something and *saying something of* it. Saying that London is pretty, for example, involves referring to London and saying of it that it is pretty. One way of doing this is to use 'London' to refer to London and 'is pretty' to say of it that it is pretty. Following ordinary usage, the *speaker* is taken as the understood subject, so to speak, of 'refer to' and 'say of'. Initially, we should deplore the unannounced slipping and sliding, back and forth, between different subjects of 'refer to' and 'say of', and we should be suspicious of talk about what *uses of words* refer to and say of things, and of talk about what *words-relative-to-contexts* refer to and say of things—unless such talk is taken to be straightforwardly translatable into talk about things that speakers are doing. But we should be open to the idea that new, *technical* uses of 'refer to' and 'say of' may emerge in the course of our inquiries. Who or what *A* is referring to by uttering some expression *X* is determined by *A's referential intentions* in uttering *X*. (Talk of 'contexts' 'fixing' or 'determining' the references of expressions—for example the references of indexical expressions—must involve some form of confusion.)

Nonetheless, a distinction between what *A* referred to and what *A intended to refer to* is not one obviously lacking a point. In the first instance we should separate (i) *who or what A intended to refer to by an expression X on a given occasion*, and (ii) *who or what a rational, reasonably well-informed interpreter in B's shoes thinks A intended to refer to by X on that occasion*. In cases where (i) = (ii), we can talk freely about *what the speaker referred to*. In cases where (i) ≠ (ii), we *could* argue about which of (i) or (ii) or some third thing has the 'right' to be called *the person of thing referred to*, but what would be the point? First, what third thing distinct from (i) and (ii) could be of any significance to a theory of

interpretation? There is simply no rôle for a transcendent notion of what was referred to upon which (i) and (ii) converge when all goes well. Referring is like saying and implying: it is the *coincidence itself* of (i) and (ii) that constitutes success, and it is the *potential* for such coincidence, independently of some third thing, that *gives sense* to the very idea of referring. (Contrary to linguistic appearances the concept of *intending* to refer to  $X$  is, in fact, more basic than the seemingly simpler concept of referring to  $X$ .) Second, why is a choice between (i) and (ii) needed in cases where (i)  $\neq$  (ii)? Conceptually they are distinct, and they are both needed in a theory of interpretation. When all goes well, they coincide, and it's just too bad they don't always do so. There is no philosophical payoff in bestowing the honorific 'what was referred to' on one rather than the other, or on some third thing, when they diverge. Referential intentions are constrained by belief and knowledge. Assuming he is being co-operative,  $A$  cannot (intend to) refer to some particular person  $\alpha$  by uttering some expression  $X$  on a given occasion if he believes it is impossible for his audience  $B$  (or at least any rational, reasonably well-informed interpreter in  $B$ 's shoes) to construe him as referring to  $\alpha$ .<sup>31</sup> Among the things that constrain the formation of  $A$ 's referential intentions are  $A$ 's knowledge of the meanings of the referring expressions he is using and his (tacit) knowledge of the syntax of the language he is using (which may bear on the matter of co-reference or binding where pronouns are concerned).<sup>32</sup>

12. *Aphonicity*. It is now time to get more theoretical. A distinction between PF ('Phonetic Form') and LF ('Logical Form') in something like Chomsky's sense, is almost certain to play a key rôle in a theory of interpretation, where a sentence's PF is (roughly) a representation that expresses its phonology, and its LF a representation that expresses *all* syntactic properties relevant to interpretation.<sup>33</sup> This distinction brings with it the possibility of revealing in the LF of a sentence  $X$  syntactic objects that have no counterparts in  $X$ 's PF. Such aphonic (phonologically null) expressions are as much in need of interpretation when  $X$  is uttered as any other elements in  $X$ 's LF. (If a sentence's LF expresses *only*

---

<sup>31</sup> I mention *co-operation* because in certain circumstances one may seek to disguise one's intended referent from others, for example in cryptic poetry, diary entries, or dramatic irony. In such cases, there is either no intended audience distinct from oneself or some individual distinct from oneself with whom one is engaged but with whom one is being less than fully co-operative in the sense discussed earlier. The issues here are intimately connected to Grice's (1989) discussions of communicative intentions in the absence of an audience and to the issue of whether one's future self constitutes an audience.

<sup>32</sup> As it turns out, I think there are no syntactic constraints on co-reference only on binding. See below.

<sup>33</sup> See Chomsky (1986, 1995, 2000).

syntactic properties relevant to interpretation, as current theory dictates, this becomes a matter of definition.)

13. *Indexicality*. Identifying the LF of a sentence *X* does not constitute identifying what *A* says on a given occasion by uttering *X*. For one thing, *X*'s LF may contain an indexical expression like 'I' or 'he' or 'that'. So identifying *X*'s LF still leaves *B* some interpretive work to do, work that will involve accessing and integrating all sorts of information not carried or revealed by the LF itself.<sup>34</sup>

14. *Anchoring*. Idealization and abstraction from the details of particular speech situations or contexts are unavoidable if work is to proceed. To this extent, we may temporarily avail ourselves of the formal 'indices' or 'contexts' of indexical logics in order to anchor or co-anchor the interpretations of indexical or anaphoric expressions that are not of primary concern at a certain point of investigation. We should not take formal indices themselves particularly seriously, however. They are useful transitory tools, methodological or heuristic devices, not serious posits in a theory of utterance interpretation.

15. *Mongrels*. Since LFs may contain aphonics and may contain indexicals, we should be open to the possibility that they may contain aphonic indexicals. (At the same time, we should no more take seriously the idea that pairing a formal 'context' with an aphonic indexical in a sentence *X eo ipso* solves a genuine problem about the *interpretation* of an utterance of *X* than we should take seriously the idea that pairing a formal 'context' with a *phonic* indexical does so.) Aphonic indexicals are not the only possible mongrels. Since LFs may contain aphonics and may contain bindable variables, we should be open to the possibility that they may contain aphonic bindable variables. (Compare 'everyone wants John to leave' and 'everyone wants to leave'. Perhaps the subject of 'to leave' in the latter is an aphonic variable bound by 'everyone'. Certainly it is not an aphonic copy of 'everyone'.) And why not aphonic, indexical, bindable variables?

16. *Isomorphism*. It is at least methodologically useful to say that identifying what a speaker said by uttering a sentence *X* on a given occasion involves 'entertaining' a sentence of Mentalese; or that it involves or 'entertaining' a

---

<sup>34</sup> This does not mean that the pragmatist cannot, for certain expository or investigative purposes, operate *as if* a description of a sentence's LF gives us a description of what *A* said *relative to certain heuristic stipulations*. Formal 'contexts' or 'indices' are used in logic to anchor or co-anchor indexical elements in order to cancel or pair their effects across similar structures. Without commitment to the view that formal contexts play any sort of rôle in a theory of interpretation, the linguistic pragmatist may sometimes borrow this technique in order that a particular investigation may focus on particular non-indexical properties of LFs that are relevant to a theory of utterance interpretation.

structured proposition. (Perhaps entertaining a sentence of Mentalese ultimately amounts to entertaining a structured proposition because it involves entertaining a representation whose rôle in our mental life can be explained only in terms of it having a certain ‘content’ that a structured proposition supplies. Who knows?) With Sperber and Wilson, let us work for the moment with Mentalese. Since LFs are not full-blown representations of Mentalese (and so do not express propositions), but only ‘blueprints’, ‘schemas’, ‘skeletons’, or ‘templates’ for such, in advance of serious empirical investigation we cannot rule out the possibility of a *failure of isomorphism* in the mapping between the LF of a sentence *X* and a Mentalese representation the entertaining of which constitutes understanding what *A* said by uttering *X* on a particular occasion (or in the mapping between *X*’s LF and a structured proposition, the entertaining of which constitutes understanding what *A* said by uttering *X* on a particular occasion). That is, we cannot rule out atoms of the Mentalese representation (or atoms of the structured proposition) to which no element of *X*’s LF corresponds. (*A* might utter the sentence ‘The embassy is closed’ on a particular occasion and *B* may be required to entertain the Mentalese sentence all too conveniently rendered as THE U.S. EMBASSY IN LONDON IS CLOSED in order to grasp what *A* said.<sup>35</sup> *A* may utter, ‘the ham sandwich wants extra pickles’, and *B* may be required to entertain the Mentalese sentence THE MAN WHO JUST ORDERED A HAM SANDWICH WANTS EXTRA PICKLES. *A* may utter, ‘the hostages landed back on American soil today’ and *B* may be required to entertain the Mentalese sentence THE FORMER AMERICAN HOSTAGES AT THE U.S. EMBASSY IN TEHRAN LANDED BACK ON AMERICAN SOIL TODAY. It would seem that *A* may even utter less than a whole sentence—for example, ‘no thank you’ or ‘a cappuccino, please’—and thereby say something.

We cannot rule out the possibility, however, that future work in syntax will indicate that we are closer to isomorphism than superficial appearances suggest, for all sorts of aphonics in LF may be revealed. Presumably, Mentalese representations will have to contain elements that function as (or at least do the work done by) bound variables, so we may well have to consider the possibility that interpreting a particular utterance of a sentence *X* may involve entertaining a Mentalese representation that contains a mental variable with no counterpart in *X*’s LF. On the other hand, syntactic evidence might be found for the existence of an aphonic variable in *X*’s LF. We cannot dogmatically assume that there *must* be isomorphism, and we should recoil from the unargued goal of attaining

---

<sup>35</sup> If you know English and your shift-key works, Mentalese is a cinch (Mentalease?). Structured-Propositionese is a little harder: you need good angled brackets.

isomorphism by freely adding aphonics to LFs as if adorning some garish Christmas tree with a new light wherever it seems too dark. (We shouldn't get hooked on aphonics.)

17. *Ellipsis*. Corresponding to the sentence-utterance distinction impressed upon us so forcefully by Austin, Grice, and Strawson, we must take seriously two important and distinct uses of the words 'ellipsis' and 'elliptical.' The first is a strict linguistic (or grammatical) notion found in talk of elliptical *sentences* in generative linguistics, a notion sometimes called *deletion* and which involves erasing elements in the generation of PF representations.<sup>36</sup> Linguistic ellipsis concerns the superficial incompleteness of *structures*, and as such is subject to a stringent condition on the constancy of form and interpretation that has been investigated by linguists under the rubric of *recoverability*. (*A* can use the sentence 'I can tango but Mary can't' to say that he can tango but Mary can't *tango*, but not to say that he can tango but Mary can't *sing*. This is because it is elliptical for the complete sentence 'I can tango but Mary can't tango.') The second notion of ellipsis is a pragmatic (or speech act) notion, found in talk of elliptical *utterances* of (elliptical on non-elliptical) sentences.<sup>37</sup> Pragmatic ellipsis concerns the incompleteness of *interpretations*, and as such is governed only by general pragmatic principles of interpretation. (*A* can use the sentence 'I'm going to a party at the embassy' to say that he's going to a party at the British embassy in Athens, for example, or to say that he's going to a party at the US embassy in London because there is no particular complete sentence that the sentence *A* uttered is an incomplete or elliptical version of.) We must accept that people often speak elliptically without much (if any) conscious effort and that hearers interpret elliptical utterances without much (if any) conscious effort. In such situations typically the speaker and hearer can both readily expand upon the sentence uttered in such a way that explains the ellipsis. Many distinct expansions may be perfectly acceptable, what is said being to a degree indeterminate. Many distinct expansions may be perfectly acceptable, what is said being (to a lesser degree than what is implied) somewhat indeterminate

18. *Competence*. Three major components of a theory of interpretation are a *syntactic theory*, a *semantic theory* and a *pragmatic theory*. Certain preconceptions about the labels 'syntactic', 'semantic', and 'pragmatic' need to be put to one side

---

<sup>36</sup> See, for example, Williams (1977), Sag (1976), Heim and Kratzer (1998), and May (2002). The last of these provides a particularly clear and user-friendly discussion of linguistic ellipsis. The grammatical notion clearly has its roots in talk of ellipsis and elision in some traditional grammars.

<sup>37</sup> See, for example, Quine (1940), Sellars (1954), Bach (1981), Salmon (1982) and Neale (1990).



if the pragmatist position is to be understood, for these words are used in very precise ways. (Self-serving edicts from those who claim to have isolated the ‘correct’ way of making the semantics-pragmatics distinction or the ‘correct’ uses of the terms ‘semantics’ and ‘pragmatics’ should be ignored.) A *syntactic* theory for a person *A* who speaks a language *L* is an abstract description of *A*’s syntactic *competence* (in Chomsky’s sense), *A*’s tacit knowledge of the syntax of *L*. This only becomes interesting when we are clear about what counts as a syntactic fact or phenomenon. Are binding and scope syntactic phenomena? A *semantic theory* is an abstract description of *A*’s semantic *competence*, his knowledge, tacit or otherwise, of the semantics of *L*. This only becomes interesting when we are clear about what counts as a semantic fact or phenomenon. Are binding and scope semantic phenomena? Binding shows that a sharp division between syntax and semantics is illusory (which is why Chomsky is prepared to use the label ‘syntactic’ in connection with much of what many philosophers and linguists label ‘semantic’). Drawing a sharp line between semantics and pragmatics is straightforward. A *pragmatic* theory transcends individual speakers and particular languages. It is an abstract description of the mechanisms that make it possible for interpreters to identify what a speaker means by uttering a sentence (or sentence fragment) *X* on a given occasion given (at most) what a semantic theory has to say about *X*. As such, a pragmatic theory is a description of an intentional and richly inferential system, our common pragmatic *competence*. There is no assumption here, nor is there any antecedent reason to suspect, that the semantics-pragmatics distinction just drawn will be co-ordinate with the saying-implying distinction.

19. *Semantics*. Words and the ways in which they can be combined have properties that enter into an explanation of why speakers use the particular combinations they do, and why hearers interpret in the ways they do. The two most obvious properties are the *meanings* of words and *syntax*. Qua description of semantic competence, a semantic theory for a language will explain how the syntactic structure of a sentence (or sentence fragment) *X* and the meanings of the individual words in *X* conspire to constrain what speakers can *say* using *X*. Flushing out the modal: a semantic theory for a language *L* will provide, for each sentence *X* of *L*, a *blueprint* for (a *template*, a *schematic* or *skeletal* representation of) what someone will be taken to be saying when using *X* to say something. The blueprint associated with *X* is its *semantics*, and the set of such blueprints, one for every sentence of a language *L*, is the *semantics* for *L*. (The study of these blueprints is also called *semantics*. The study of the rôle of word meanings is called *lexical semantics*; the study of the rôle of syntax is called *compositional semantics*.) Semantic competence comprises at least (i) knowledge of the meanings

of individual words and (ii) knowledge of syntax (syntactic competence). It is a matter of debate whether it involves more. On the one hand, if *A* claims not to understand a sentence *X*, then it would seem that either the meaning of some word in *X* eludes him or else some aspect of *X*'s structure (ultimately *X*'s LF) does. On the other, having a model aeroplane kit, a foolproof set of instructions, excellent glue, plenty of space, good lighting, and the fingers of a heart surgeon is not the same thing as having the model aeroplane (that's why they write 'kit' on the box). Settling this debate involves settling (among other things) what syntactic competence amounts to and how bad the model aeroplane analogy is. (Certainly knowledge (in the requisite sense) of syntax does not amount to the propositional representation of a set of syntactic rules).

20. *Pragmatics*. Whereas each language (perhaps even each idiolect) has its own syntax and its own semantics—which is not to say that vital syntactic and (hence) semantic properties are not shared across languages as a result of our common biological endowment—there is, so to speak, only *one* pragmatics. *Qua* description of our shared pragmatic competence, a pragmatic theory will explain how interpreters identify what a speaker means by uttering a sentence (or sentence fragment) *X* on a given occasion given (at most) what a semantic theory has to say about *X*.

The semantics-pragmatics distinction, thus construed, is not co-ordinate with the saying-implying distinction. What *A* means by uttering *X* on a given occasion comprises what *A* *said* and what *A* *implied*. So a pragmatic theory will explain how interpreters identify what *A* said and implied by uttering *X* on that occasion given (at most) what a semantic theory has to say about *X*. If a *pragmatic* theory explained only how interpreters identify what *A* *implied* given (at most) what the speaker *said* as 'input', a gaping hole in our taxonomy of *theories* would appear. A *semantic* theory specifies the constraints that word meanings and syntax place on what *A* can *say* by uttering *X*, a blueprint for *X*. What would we call a theory that explains how interpreters identify what *A* *said* on that occasion? Not a *semantic* theory, for that specifies only a blueprint for what *A* *said*, i.e. the *sort of thing* he said. Clearly, a pragmatic theory has *two* rôles in a theory of interpretation. Even if an utterance of a sentence *X* always wore on its sleeve an unambiguous representation of its syntactic structure with no ambiguous elements, a semantic theory could still fail to identify fully what *A* said by uttering *X* on a particular occasion. For one thing, *X* may contain 'he', 'this', 'here' or 'John', in which case

the interpreter needs to identify who or what *A* is referring to.<sup>38</sup> Since these words are not ambiguous in the way /pen/ or /bank/ are said to be, and since each is not merely the unambiguous surface form of a context-insensitive definite description, something other than a semantic theory must be invoked.<sup>39</sup>

---

<sup>38</sup> My wording should make it clear that I am putting aside, for now, the fact, stressed by many pragmatists, that a pragmatic theory may have to be invoked in order to identify what *sentence A* uttered because of ambiguities at PF. Among the things a hearer or reader has to do in order to identify what *A* is saying on a given occasion, is identify which *words A* is using. /bank/ is the superficial form of either a single, ambiguous word of English or else of two distinct unambiguous words, and I do not want one's position on this matter to impinge upon one's understanding of the Insufficiency Thesis (see point 21). If /bank/ is the superficial form of a single, ambiguous word, then identifying what *A* is saying when he utters, 'I'm going to the bank' involves identifying which *meaning A* has in mind for /bank/; if /bank/ is the superficial form of two distinct, unambiguous words then identifying what *A* is saying when he utters /I'm going to the bank/ involves identifying which of the two *words A* is using. The latter view seems more useful in theorizing about language. 'Word' and 'sentence' are quasi-technical terms, there are no ambiguous words or sentences, and (following Chomsky) every sentence comprises a superficial form PF and an underlying form LF, the former being what is relevant to speech perception, the latter what is relevant to speech comprehension. When I wish to talk explicitly about an expression's PF or about the sound common to two expressions, and when I wish to avoid commitment one way or the other as to whether I am talking about one expression or two, I shall borrow the old slash notation of phonology (but with standard orthography rather than a phonological representation enclosed, as in /bank/) to individuate coarsely in terms of phonological properties. (On one use, then, /bank/ is what Perry (1998) calls a *vocable*.) Thus I sometimes use /he/, /him/, and /his/ because it is arguable that each corresponds to two distinct words in English, one that is bound and another that is not, an idea that appears to gain some plausibility when we consider, say, Scandinavian languages where something close to the distinction appears to be lexicalized. Similarly, I sometimes use /the/ so as not to prejudge the issue on the matter of a purported ambiguity in the definite article(s). Of course if /he/ and /the/ really are ambiguous, the ambiguity in question will have to be more systematic than the sort found with /pen/ or /bank/. It is easy enough to cause trouble for my use of the slash notation. Almost certainly we want to distinguish the phonologically identical but orthographically distinct 'so', 'sew' and 'sow<sup>a</sup>' (as in *seeds*), distinguish the orthographically identical but phonologically distinct 'sow<sup>a</sup>' (as in *seeds*) and 'sow<sup>b</sup>' (as in *pig*), and distinguish the phonologically and orthographically identical 'pen<sup>a</sup>' (as in *writing instrument*) and 'pen<sup>b</sup>' (as in *enclosure*); but probably we shall not need not bother distinguishing the (merely) orthographically distinct 'judgment' and 'judgement' or the (merely) phonologically distinct 'cóntroversy' and 'contróversy.' (Actually, I'm not so sure I should have said that: perhaps there are worries here not entirely unconnected to those Kripke (1979) brings up in connection with 'Paderewski'.) So when I want to individuate coarsely in terms of (roughly) phonology, I use the slash notation. Thus /pen/ and /sew/ (and, unfortunately, /sow<sup>a</sup>/ and /sow<sup>b</sup>/). If a single word can have two distinct orthographies ('judgment' and 'judgement') and a single word can have two distinct phonologies ('cóntroversy' and 'contróversy), should we explore the idea that a single word can have two distinct orthographies *and* two distinct phonologies? Or should we say that something is a single word only if it is grounded in a single phonology or a single orthography? Or should some intermediate position be explored that invokes etymology or the relative similarity of distinct orthographies and phonologies, a position according to which /doctor/ and /physician/ would be too far apart to qualify, orthographically, phonologically, and etymologically? (Notice how the notation just exploded.) Many Greek villages or islands still have two names, and the reason we talk this way is because the names seem too far apart to count as a single name (e.g. 'Thíra' and 'Santoríni'). But what about 'Aperáthou' and 'Apeíranthos'?

<sup>39</sup> Three points must be separated here: (i) failure of lexical ambiguity, (ii) context sensitivity, and (iii) rigidity.

The slack is taken up by a *pragmatic* theory: identifying what *A* said involves the exercise of cognitive capacities that integrate the semantic information carried by the sentence uttered and all sorts of ‘pragmatic’ or ‘contextual’ information including, but not limited to, information obtained by perception from the physical environment, information about the interpretation of prior utterances in the conversation (if any), information in memory, and information about how people typically behave, particularly in communicative exchanges. That is, identifying what *A* said involves processing not only the semantic information encoded in a sentence’s form, but accessing and processing information that must be picked up by listening, watching, remembering, hypothesizing and inferring, essentially the capacities exercised in identifying what *A* implied. To this extent, then, identifying what is said is a *pragmatic* as well as a semantic matter. It involves *pragmatic inference* as well as *linguistic decoding*. Identifying what a speaker implied is something explained by a pragmatic theory, typically taking into account what *A* said; but identifying what the speaker said is also something explained by a pragmatic theory, taking into account (in a big way, to be sure) a sentence’s blueprint, which is explained by a semantic theory. Underpinning the difference between identifying what *A* said and what he implied is a distinction in the type of typical *input*: to identify what *A* said on given occasion by uttering *X* the pragmatic system typically takes as its primary input what the speaker said. This leaves many questions open: (1) To what extent is pragmatic processing deductive? (2) To what extent does it take place unconsciously? (3) What sorts of things affect its speed? (4) To what extent is it task-specific or modular?

21. *Underdetermination*. It is now possible to bring together several points. The rôle of a pragmatic theory in identifying what *A* said by uttering *X* on a given occasion is not restricted to identifying who or what *A* is referring to by any referential expressions in *X*. Saying involves referring and predicating; and just as identifying what *A* is up with any referential devices in *X* involves more than consulting a mental lexicon, so does identifying what he is up to with any predicative devices in *X*. It may, for example, require the ‘saturation’ of an ‘implicit argument’ as in ‘It’s raining’ or ‘I’ve finished’. (Some implicit arguments may be mandated by syntax as well as semantics.) Or it may require ‘enriching’ a predicate in some way that is reasonably obvious and presumably acceptable to *A*. The sentence ‘every woman has a job’ might be used to say that every woman in Flint has a job, or that every woman in Woodside has a job, etc.

It will pay to separate and rejoin two points here, one epistemological, the other metaphysical, both intimately connected to the points made earlier about *Asymmetry* and *Reciprocity*. The epistemological point concerns *insufficiency*, the

metaphysical point *underdetermination*. From the hearer's perspective, we can talk first about the fact that knowledge of the syntax of *X* and knowledge of the meanings of all the words in *X* do not suffice for identifying what *A* is saying by uttering *X* (even where the superficial form evinces no lexical or structural ambiguity). At most they yield a blueprint. Now we can bring in *A* himself. What *A* says is wholly determined by certain specific intentions he had in speaking, intentions massively constrained by his knowledge of syntax and word meaning (and a whole lot more). *A* tacitly knows that *B*'s knowledge of word meaning and syntax will not suffice to furnish *B* with a complete account of what he has said. We can now introduce some theoretical shorthand to obviate the need to keep talking about speakers' and hearers' knowledge, tacit or otherwise. Let us say that syntax and word meaning together *underdetermine* what is said (all the time remembering this is shorthand). But we are not yet where we need to be. The first thesis we need is this:

(IT) *The Insufficiency Thesis*: Identifying what a speaker or writer, *A*, is saying by uttering an unambiguous, declarative sentence *X* on given occasion goes well beyond recovering *X*'s underlying syntax, knowing the meanings of all of the words in *X*, and identifying who or what *A* is referring to by any referential expressions in *X*.

This goes beyond the insufficiency just mentioned because it entails that even when *B* has identified who *A* is referring to by any referential expressions in *X* (/John/, /he/, /here/, /that/, /I/, /you/, and so on), *B* still doesn't have everything he needs to identify what *A* said. *A* tacitly knows that *B*'s knowledge of word meaning, knowledge of syntax *and* knowledge of who or what *A* is referring to by any referring expressions in *X* will not suffice to furnish *B* with a complete account of what he has said. Corresponding to this we can formulate the shorthand we really want:

(UT) *The Underdetermination Thesis*: What *A* says by uttering an unambiguous, declarative sentence *X* on given occasion is underdetermined by *X*'s syntax, the meanings the words (and any other morphemes) in *X* (and the meanings, if any, of prosodic features of *X*), and the assignment of references to any referring expressions in *X*.<sup>40</sup>

---

<sup>40</sup> The Underdetermination Thesis is regularly stressed by linguistic pragmatists, and some view it as a cornerstone of the general outlook. This use of 'underdetermination' is found in the work of Sperber and Wilson (1986, 1995) and borrowed by many of those they have influenced, including Bezuidenhout, Blakemore, Carston, Papafragou, Récanati, and Rouchota, as well as me in *Descriptions* (1990: 105 n.16; p. 114 n. 46; p. 116 n. 54). In the language of Perry (1986, 1998, 2001), talk of underdetermination is roughly equivalent to talk of constituents of propositions expressed that are 'unarticulated', i.e. constituents corresponding to no constituents of the sentence

As far as constructing and evaluating a theory of interpretation are concerned, we must make sure we separate talk of (a) the *interpretive target* (stipulated in advance by the theorist in particular cases on the basis of intuitive judgment, modulo reflective equilibrium with the best theory up to that point), and talk of (b) the *knowledge and mechanisms* in play (under investigation and hypothesized by the theorist). And in talk of knowledge and mechanisms we must be careful separate (i) the role of syntax; (ii) the role of word meaning; and (iii) the actual pragmatic mechanisms.

An interpretive target is a characterization of what our intuitive judgments reveal the speaker to have said and which an adequate theory of interpretation should deliver. It is common to specify interpretive targets using more language: By uttering ‘she has a job’ *A* is saying that Margaret Thatcher has a job; by uttering ‘every woman has a job’, *A* is saying that every woman living in Woodside has a job. There is, of course, something a bit funny about this, for surely we can now ask for a characterization of *what the theorist said* when he uttered the sentence ‘By uttering “every woman has a job”, *A* is saying that every woman living in Woodside has a job’. Nonetheless, this is what we do, and when pressed often we wheel out some set theory: By uttering ‘every woman has a job’, *A* is saying something that is true iff at the time of utterance *t*, the intersection of the set of things that are women at *t* and the set of things living in Woodside at *t* is a subset of the set of things that have jobs at *t*. While this may provide us with the conditions under which what *A* said is true, it falls short of specifying *what A* said for familiar reasons. (First, *A* is sure to deny it. Second, surely *A* would have said something different had he uttered, ‘every woman has a job and  $19^2 = 361$ ’). So

---

uttered. (As Perry sometimes puts it, we don’t always articulate things when it’s clear from context what they are.) I don’t know why under-determination is rejected so vehemently by some philosophers of language, but I have a suspicion two related factors may be implicated, both involving fear and philosophical temperament. The first is a simple unwillingness to concede apparently hard-earned territory, a reluctance to accept that some of the traditional problems involved in so-called ‘compositional semantics’ are actually the products of specious questions in the philosophy of *language* with genuine and important counterparts in the philosophy of *mind*, mostly about *inference* and *the composition of thought* (a reluctance to accept that as far as natural language is concerned, trying build pure *content* is as futile as trying to build pure *character*). The reluctance, it seems to me, amounts to little more than obstinacy or fear of a philosophical pink-slip. The second fear might be viewed as an extension of the first: the fear that if there is not at least one component (what is said) of what a speaker means that can be nailed down precisely and completely without taking into account too many ‘pragmatic considerations’, then systematic semantics as typically understood is doomed, and with it any chance of producing a serious theory of language. This fear seems to me entirely unwarranted. Natural language semantics may not be quite as straightforward or far reaching as many have thought, but there is plenty of systematic semantics for all of us (and more) to do for longer than we will ever have to do it.

when pressed again we wheel out something like a *situation* or a *structured proposition*.

Where *knowledge* of syntax, word meaning, and the theory of blueprints are concerned, there is much work to be done by philosophers of language and linguists together. As far as *mechanisms* are concerned, we are squarely in the realm of psychology, and some philosophers of language and linguists may well opt out here. The psychological part of the overall project may certainly be informed by philosophical reflections, such as Grice's, on the nature of rational, purposive behaviour, but ultimately it is a wholly empirical enterprise the aim of which is to identify the cognitive mechanisms whereby the hearer effects the relevant identifications on a given occasion.

22. *Indeterminacy*. What a speaker says and implies may be indeterminate in at least the following sense: in any vocabulary in which what someone says or implies can be usefully specified, there will be alternative and strictly distinct specifications between which no principled choice can be made. We should not worry that indeterminacy of this sort presents problems for particular semantic proposals. (For example, we should not regard traditional accounts of descriptions as damaged in any way by the indeterminacy attaching to intuitive 'completions' of those that are said to be 'incomplete' relative to particular occasions of utterance.)

23. *Convergence*. A univocal saying-implying distinction is *empirical*, *ordinary*, and *beneficial (practical)*. The distinction is empirical insofar as it assumes that, typically, representations corresponding to what *A* said and implied are the outputs of cognitive mechanisms involved in the interpretation process. It corresponds to something entrenched in ordinary talk (despite the fact that we may disagree in particular cases). And it underpins the very idea of codifying principles meant to regulate societies and the behaviour of their members (e.g. laws, contracts, and commitments) by virtue of being a distinction one side of which (saying) is about as *objective* as anything can be, a fact itself guaranteed by the empirical and ordinary nature of the distinction. (To say this is not to say there cannot be disputes about what was said, changes of opinion after discussion of problematic cases, or specialists (or at least professionals) in societies to whom tough cases are referred when the issue needs forcing. Rather it is to say that there is enough overlap in judgment to render regulation, commitment, and so on meaningful notions.

24. *Formalism*. Advances in our thinking about language have come out of developments in logic and formal philosophy, particularly by way of the

construction and use of various types of broadly mathematical theories, systems or analyses—the predicate calculus, model theory, modal logics, set theory, recursion theory, and generalized quantifier theory to name the most obvious. But it does not follow that associating utterances with models, possible worlds, structured propositions, indices, functions or even favoured formalisms *ipso facto* constitutes part of a theory of utterance interpretation. Rigorous formalism almost certainly has its place in articulating certain parts of a theory of interpretation; but a favoured mathematical idea and an associated formalism must not so dominate our inquiry that the questions motivating it in the first place become obscured or transmogrified to the point of demanding purely technical answers. We should strive to use our formalisms judiciously, sparingly, only where they were needed to effect a useful idealization or abstraction, forestall a potential ambiguity, capture a generalization, facilitate a transition, or usefully abbreviate something. Appeals to, say, higher-order functions or set-theoretic entities (and the use of corresponding notations) are ultimately dispensable in the theory of interpretation, sets and functions being no more than occasional, transitory tools of no intrinsic interest outside mathematics proper and the philosophy thereof. Such entities are not the objects of semantic investigation themselves, and it should not be a goal of any branch of philosophy to drag them into investigations whenever the opportunity presents itself.

#### 4. Pronouns and Anchors

Following Grice, let us talk about what a speaker *A* means on a given occasion by uttering some sentence *X*, factoring this into what *A* says and what he (merely) implies. As Grice notes, *identifying* what the speaker is saying is not simply a matter of identifying *X* and recovering its linguistic meaning (blueprint), if only because of the existence of pronouns. Unlike some of his critics, Grice is careful not to run together epistemological and metaphysical points here, despite their evident interconnectedness. The important metaphysical question is: what *determines* what a speaker said on a given occasion? And the Gricean answer is: certain specific *intentions* he had in producing his utterance, intentions that are severely constrained by his tacit grasp of syntax, of the meanings of the words he uses, and of the way rational, co-operative beings function, his beliefs about the audience, about the context, and about the topic of conversation, and probably a whole lot more.<sup>41</sup> The important epistemological question is: what knowledge or

---

<sup>41</sup> See Grice (1989).



information does a hearer use in *identifying* what the speaker said? And the Gricean answer is: by his tacit grasp of syntax, of the meanings of the words used, and of the way rational, co-operative beings function, his beliefs about the speaker, about the context, and about the topic of conversation, and just about anything else he can get his hands on.

Let us take a concrete example. If I say something by uttering a sentence *X* that contains the personal pronoun ‘he’ and the demonstrative ‘this’, then (in the simplest case, at any rate) my referential intentions determine who I mean by ‘he’ and what I meant by ‘this’. (Similarly, my lexical intentions determine what I meant by /bank/ if *X* contains one of the words we write that way.) Your job as hearer is to identify what I meant by uttering *X*, and very likely you will not succeed unless you identify who I meant by ‘he’ and what I meant by ‘this’. (Similarly, what I meant by /bank/.)

Sperber and Wilson (1986, 1995) point out that pronouns are just the tip of a pragmatic iceberg for traditional, titanic accounts of what is said: quite generally, what a speaker says is underdetermined by the meaning of the sentence uttered, even relative to reference assignment. Now one can perfectly consistently accept Sperber and Wilson’s underdetermination thesis, as Récanati (1987, 1989, 1993, 2001) does and as I do in *Descriptions* (1990: 105 n.16; 114 n. 46; 116 n. 54), without rushing to embrace the details of their Relevance Theory. For that theory is meant to provide an account of the mechanics of utterance interpretation, of the richly inferential processes providing the basis of an empirically satisfying account of how interpreters (i) identify which sentence a speaker has produced on a given occasion in cases where identification of phonological form fails to yield a unique result; (ii) identify what the speaker *said* by uttering *X* on a given occasion in cases where identification of the meaning of *X* falls short; and (iii) identify what the speaker implied by uttering *X* on that occasion. Relevance Theory goes well beyond the thesis that the meaning of a sentence *X* may underdetermine what a speaker says by uttering it on a given occasion and well beyond the vague Gricean idea that quite general principles governing the way we reason about the behaviour of others lie at the heart of an explanation of how we communicate.

In principle just about any information could be relevant or brought to bear on interpretation, and one of the main problems involved in constructing a pragmatic theory is explaining how the information that actually *is* brought to bear is delimited.<sup>42</sup> The second problem concerns *how* it is brought to bear. Linguistic

---

<sup>42</sup> There is a division here between the optimists and the pessimists. Unlike Sperber and Wilson, Blakemore (1987), Carston (2002), and other Relevance theorists, Fodor (1983, 2001), Chomsky

pragmatism finds little sense in the idea that two quite distinct sets of information-gathering and inferential mechanisms are at work when a hearer tries to identify what a speaker means, one set that works on sentence meanings and yields what the speaker said, and another set that works on what the speaker said and yields what he meant but did not say (i.e. what he implied).

It is odd that some philosophers write as if (or even claim that) two quite distinct sets of cognitive mechanisms must be at work. I detect two related ideas lurking behind this assumption: (i) something to do with ‘simplicity’ (or ‘degree of difficulty’) or ‘systematicity’ (or ‘range of possibilities’); (ii) the influence of ‘indexical logics.’

(i) Whilst it is true that identifying who or what a speaker intends to be referring to on a given occasion by some particular referring expression *X* is constrained by the linguistic conventions governing the use of *X*, this does not necessarily make matters particularly straightforward or reduce the number of hypotheses that could, in principle, be investigated and assessed. Consider the interpretation of an utterance of the pronoun ‘it’. As Sperber and Wilson (1986: 187) note, all that the linguistic conventions governing the pronoun ‘it’ insist upon, in any context, is that the object should be non-human, giving every hearer in every context an indefinitely large choice of possible referents. And surely the same general considerations about, say, relevance, truthfulness, informativeness or whatever, that are invoked in identifying what a speaker is implying on a given occasion will be invoked in identifying who or what a speaker is referring to by ‘it’ on a given occasion.

The point can also be made in connection with an incomplete description like ‘the table’ or some other incomplete quantifier expression like ‘every man’ or ‘no one’. It is sometimes said that identifying what *A* says by uttering a sentence containing such an expression involves either (a) coming up with an appropriate domain of quantification *implicit* in the utterance (the ‘implicit’ approach), or (b) coming up with an appropriately ‘richer’ nominal *A* could have used to make his meaning more *explicit* (the ‘explicit’ approach).<sup>43</sup> Whatever the final merits of such suggestions, one thing is quite clear: the same general considerations about, say, relevance, truthfulness, informativeness or whatever, that are invoked in

---

(2000), and Davidson (1986) suspect that producing an overarching theory of interpretation will require nothing short of a complete theory of mind. For discussion, see Carston (2002).

<sup>43</sup> For detailed discussion, see Neale (forthcoming, *a*)

identifying what a speaker is implying on a given occasion will have to be invoked in identifying an appropriate domain or an appropriate completing expression.<sup>44</sup>

(ii) Many philosophers write as if (or even argue that) understanding what a speaker *A* said on a given occasion by uttering a sentence *X* with its conventional meaning is a matter determined by the meaning of that sentence and a ‘context’, in a sense of this frequently invoked word that is meant to make it more than simply a label for whatever it is that ‘bridges the gap’ between the meaning of *X* and what *A* said by uttering *X* on that occasion. For example, it is frequently claimed that all one needs to bridge the gap is some sort of formal object, an ‘index’ or ‘context’ in the form of an ordered *n*-tuple that secures the references of a few annoying ‘indexical’ pronouns (‘I’, ‘you’, and ‘he’, for example) and one or two other ‘indexical’ words that have a somewhat pronominal nature (‘here’ and ‘now’, for example).<sup>45</sup>

This idea is rightly spurned by Sperber and Wilson (in *Relevance*), by Evans (in *The Varieties of Reference* and *Collected Papers*) and by Chomsky (in *Reflections on Language* and practically every book or article he has written in which the interpretation of pronouns is discussed). For there is an implicit recognition in these works, and in many others that bear their influence, that whilst formal contexts may have a useful *methodological* rôle from time to time, they are strictly irrelevant to a proper theory of utterance interpretation.

For various semantic and syntactic purposes, it is often desirable—if not mandatory—to abstract or idealize away from facts to do with particular speech situations—‘pragmatic’ or ‘contextual’ factors, as they are sometimes called—in order to get on with a particular piece of work. And as long as caution is exercised there is no harm in this. For example, *with certain restricted purposes in mind*—and without any sort of absurd commitment to the idea that such entities play a role in utterance interpretation—formal ‘indices’ can be introduced to serve as ‘contexts’ with which sentences can be paired in order to ‘anchor’ or ‘co-anchor’

---

<sup>44</sup> Postulating an aponic indexical, *domain* variable in underlying syntax makes no more of a contribution to explaining *how* hearers interpret utterances than does postulating an aponic indexical *assertion* variable in underlying syntax (or an aponic indexical *irony* variable). That is, the interpretive task facing the hearer is made no easier by the existence of an aponic contextual variable, even when, as in the case of the supposed assertion or irony variables, there are just two possible values to choose from. And of course, giving phonetic form to such operators—‘asserting-or-not, it’s Tuesday’ or ‘being-ironical-or-not, it’s a lovely day’—doesn’t help the hearer either. However, if it can be demonstrated that such variables are bindable, then their existence might be justified—though far from required—on minimalist assumptions (On the purported bindability of domain variables, see Stanley and Szabó (2000) and Stanley (2000, 2002).)

<sup>45</sup> The word ‘indexical’ is itself part of the problem, suggesting as it does that interpreting such devices involves merely looking something up in an ‘index’. People can be more influenced by labels than they sometimes realise.

the interpretations of certain indexical expressions. The usual idea is to construe such expressions as free variables and treat indices as sequences or functions that assign these variables values. Famously, this idea has been used to capture model-theoretically the validity of inferences whose premises and conclusions are stated using indexical sentences:<sup>46</sup>

A: If the next left is not Bank Street, that man gave you the wrong directions.

B: It's not Bank Street; so he gave me the wrong directions.<sup>47</sup>

It is paramount in such work to keep things tightly under control in the following sense: the logician wants a mechanism that can (a) scan a set of sentences for occurrences of symbols on some pre-existing list of devices that do not carry their values with them, then (b) use an index to assign a value to each occurrence of such a symbol. If this goes well, logical deductions can proceed (assuming a semantics for items of a pre-selected 'logical' vocabulary of course). If there is still slippage after the index has made its assignments, on standard assumptions there is only one solution: posit further indexical symbols in the sentences involved, symbols which are invisible in surface syntax yet revealed by an analysis of their 'logical forms', then try again.<sup>48</sup>

In the philosophy of language, indices have a methodological rôle for they can be used to anchor or co-anchor indexical and anaphoric expressions and so allow

---

<sup>46</sup> See (e.g.) Kaplan (1989), Lewis (1972), Montague (1974). A lot in this area turns on one's conception of logic, and my wording evinces a particular stance, though not one I want to insist on: logical relations hold among what is expressed by sentences not among sentences themselves. (Various issues about the notion of formal validity and inference rule must be faced (but usually are not) by people who hold this view of logic.) The point I am making in the text is not dependent upon this stance. Cf. discussions of the difference between the logical form of a proposition and the logical form of a particular sentence used to express that proposition.

<sup>47</sup> A related point might be made in connection with anaphora:

Every man loves his mother

John is a man

---

John loves his mother.

<sup>48</sup> It is, perhaps, tacit recognition of this fact that has led some philosophers to conclude that there is no hope of producing a theory of *utterance interpretation* without positing all sorts of aponic, indexical elements in the underlying syntax of natural language sentences. We may use anything we like to throw light on the syntax of natural language, but we must never lose sight of the fact that discerning the syntactic structures of our sentences is an empirical exercise. Certainly the idea of aponic elements in syntax is not objectionable in itself. On the assumption that syntax relates sound and meaning, we must certainly allow for the possibility of elements that have sound but no meaning ('it' in 'it's raining?'), or meaning but no sound (the understood subject of 'leave' in 'Tom wants to leave?'). And there can be little doubt today that great advances in our understanding of syntax have been made by those such as Chomsky who have not shied away from the idea of aponic items in syntax and argued for their existence and explanatory value. But we cannot simply *assume* that whenever we encounter some feature of what is said that does not appear to correspond to any element or feature of the sentence uttered it follows that there is some element in underlying syntax waiting to be exposed.

work to proceed more easily on *other* expressions and on what people say (and imply for that matter) by uttering them on given occasions.<sup>49</sup> However, there is an idea that has emerged from work on indexical logics for which we can have little sympathy. This is the idea that sentence meanings and contexts can be paired to provide something of *empirical* significance: what a *sentence X says relative to a context C*.<sup>50</sup> We must not lose sight of certain facts. First, as far as utterance interpretation is concerned, such ‘contexts’ are strictly irrelevant. Utterances do not come with such devices attached that anchor or co-anchor indexical, demonstrative, or anaphoric pronouns. The hearer has plenty of *pragmatic* work to do, much of it rightly called inferential, albeit inferential in a way that is steered by the meanings of individual words. A few passages from Evans (1982, 1985) summarize the situation well:

All that the conventions governing the referring expression ‘he’ insist upon, in any given context, is that the object referred to should be male. (1982: 312) There is no linguistic rule which determines that a ‘he’ or a ‘that man’ refers to *x* rather than *y* in the vicinity, or that it refers to someone who has just left rather than someone who has been recently mentioned (1985: 230-1). ‘This’ and ‘that’ are even less specific, contributing merely the vaguest suggestion of a contrast between nearer and further (in some generalised sense). . . [Footnote: Often the *predicate* does more to narrow down the range of possible interpretations of the referring expression than does the referring expression itself . . . ] (1982: 312). Let me take another example: the expression ‘you’: If a speaker addresses a remark to someone, saying, ‘You are a

---

<sup>49</sup> See *Descriptions*, Ch. 3, for example.

<sup>50</sup> I am putting aside here some very real concerns about talk of *sentences* saying things relative to contexts. I am sceptical about the value or relevance of the use of the verb ‘say’ assumed in this way of talking to the project of constructing a theory of utterance interpretation, unless it is understood as a stylistic variant of talk of *speakers* saying things by uttering sentences on given occasions. Judgments about what a *speaker* said, and about whether what he said was true or false in specified situations, constitute the primary data for a theory of interpretation, the data it is the business of such a theory to explain. What a speaker *says* and what he *implies* (e.g. conversationally implicates) on a given occasion are the things that together constitute what the speaker *means*, and a theory of interpretation is meant to explain the role of linguistic meaning and inference in the hearer’s identification of what the speaker meant. No-one has intuitions about *what is said by a sentence X relative to a context C* or about the truth or falsity of *X* relative to *C* unless this is just a formal way of talking about what the speaker said by uttering *X* on a particular occasion—the occasion that *C* is being used to partially model. If such talk is straightforwardly transposable into talk about what the speaker said then we can accept its empirical significance. If it is not so transposable, then its empirical significance must be justified in some other way, from *within* the theory of interpretation by reference to some empirical role it is required to play in an explanation of what a speaker says and implies by uttering *X* on a given occasion, in much the same way that notions such as LF (‘Logical Form’), scope, and binding are motivated from within. If some such motivation is forthcoming, we should be only too happy to listen. I suspect it will not be forthcoming because the notion of what a *sentence says* relative to a context is going to be too thin and overly-detached from speakers’ communicative intentions to carry any empirical weight. Nonetheless, I adopt a wait-and-see approach. We are involved in an empirical enterprise after all.

crook', it is surely clear that an identification is called for on the part of the audience: in order to understand the remark, it is not enough to know that there is one, and only one, person whom the speaker is addressing, and that the speaker is saying of that person that he is a crook . . . a quite specific *kind* of identification is called for; the person addressed has not understood the remark unless he realizes that the speaker is saying that *he* is a crook. . . understanding the remark requires the hearer to know *of* an individual that he is being addressed. (1982: 314).

Nothing about the meaning of the word 'you' tells you that you are being addressed.<sup>51</sup>

Quite generally, there is something artificial about construing the meanings of (e.g.) 'I', 'we', 'you', 'he', 'she', 'it', 'they', 'this', 'that', 'these', 'those', 'here', 'now', 'there', 'then', 'today', 'yesterday', and 'tomorrow' as functions from contexts to references. The meanings of these devices, as Evans (1982, 1985), and Sperber and Wilson (1986) stress, are just *constraints* on references, more precisely constraints on the referential intentions with which the devices can be used. I am mystified why much mainstream philosophy of language misses this. We need to distinguish two ideas about formal contexts, one sensible, the other silly. The silly idea is that utterances come with pre-packaged 'contexts' that provide values for indexical expressions. The sensible idea is what I call *methodological anchoring* (*anchoring* for short). For various pragmatic, semantic and syntactic purposes, it is often helpful, perhaps even mandatory, for a theorist to abstract from various 'contextual effects' or 'pragmatic factors' in order to get on with a piece of work, and so it is sometimes useful to use an 'index' as a way of anchoring the interpretations of indexical expressions that are not, at that moment, the objects of primary concern, *even though the theorist knows the interpretation of these indexicals is not as straightforward as invoking an index might suggest*. If one is working on definite descriptions, for example, one might want to prescind, as much as possible, from the effects of, say, indexical pronouns occurring inside nominals; and if one is working on 'and', for example, one might want to prescind, as much as possible, from the effects of, say, indexical pronouns occurring inside conjuncts:<sup>52</sup>

- (1) *He* drove home and *he* drank those six beers *you* bought *him*

<sup>51</sup> As soon as we introduce anaphoric pronouns—those that are linked in some interpretive fashion to other expressions (their 'antecedents')—matters become more complicated. The reflexive 'himself' *must be* so-linked; the non-reflexives 'he', 'him', and 'his' *can be* so-linked (under certain conditions). Very roughly, reflexives cannot be 'too far away' from their antecedents, and non-reflexives cannot be 'too close' to them, putting the two in virtual complementary distribution as far as interpretive dependence is concerned, as suggested by Chomsky's (1981, 1986) Binding Theory. See below.

<sup>52</sup> The following examples are due to Deirdre Wilson.

(2) *He* drank those six beers *you* bought *him* and *he* drove home.

To this end, we might use an index to anchor or co-anchor these expressions, to keep their special features and the complexities they introduce out of the picture as it were.<sup>53</sup>

A certain amount of care is needed in the use of the word ‘semantic’ when indices are used to anchor (or co-anchor) indexical expressions. To the extent that we are investigating the conventions governing a word whose rôle cannot be set out clearly without taking into account the conventions governing other expression(s) with which it combines to form larger expressions, we may find it convenient to talk about the (derived) conventions governing the larger phrases with respect to a particular index. For example, if the semantics of ‘the’, is being investigated, it may be useful, even mandatory, to anchor indexicals so that *other* contextual effects may be monitored. And although we may want to talk about the ‘linguistic meaning’ of, the ‘semantics’ of, or the ‘conventions governing’ an indexical or any other expression, we may also wish to talk about its ‘semantic value’ relative to a particular index, the object conveniently assigned to it by an index in order that work on pressing matters is not held up needlessly.<sup>54</sup> There is no harm in such talk as long as everyone is clear about what is going on. ‘Semantic values’ in this sense, are just *stipulated interpretations*, and the anchoring it involves is quite consistent with the idea that the interpretation of indexical expressions is basically a pragmatic matter only steered by semantic constraints.

Although I have not seen the point discussed explicitly in the literature, I get the impression some ‘anti-pragmatist’ sentiment may have as its underlying source the worry that it involves self-suspension, a willingness to abstract from contextual effects in ways that are self-defeating or paradoxical. In reality, the situation is not that different from Neurath’s. The pragmatist certainly has to tread carefully, all the while monitoring for and then abstracting from aspects of what is said that are not fixed by syntax and word meaning, and as a matter of working practice, subtle

---

<sup>53</sup> Carston (1988, 1993, 2002) implicitly anchors in her examinations of ‘pragmatic enrichments’ in connection with utterances of conjunctions (indeed, it is what she implicitly does throughout). Similarly, Evans implicitly anchors in *The Varieties of Reference* (and elsewhere), Sperber and Wilson do it throughout *Relevance* (and elsewhere) and I do it explicitly in ch 3 of *Descriptions* in connection with the effects of indexicals appearing in definite descriptions such as the following:

- (i) the first person *I* saw this morning
- (ii) *my* mother
- (iii) the *present* king of France
- (iv) the girl who made *this*.

Chomsky also does something analogous to anchoring in every work in which he discusses pronouns. (I say ‘analogous’ because of Chomsky’s concerns about reference).

<sup>54</sup> This convenience is employed time and again in *Descriptions*.

and silent measures are usually taken to prevent things becoming unmanageable, measures that certainly narrow the pragmatist's options on the vexed matter of the relation between linguistic structure and the structure of thought.<sup>55</sup> Pragmatist abstractions from context are always going to be juggling acts, the artistry of which is rather like that involved in solving for several variables at once whilst looking for an unknown number of others that are not yet in the equation and cannot be located without extremely good approximate values for those that are. On the basis of perceived use, intuition, discussions with friends, books we have read and who knows what else, we isolate what we take to be the 'linguistic meaning' or 'semantics' of an expression  $\alpha$ —its invariant role in determining what someone says by uttering sentences containing it—not unreasonably confident that certain things speakers mean when they use  $\alpha$  in their speech and writing are explicable in very general terms as things they only imply, and are not of a gravity sufficient to make us question the meaning we think we have isolated. Holding the meaning of  $\alpha$  constant, we go on to investigate and isolate the meaning of  $\beta$  and find we can make some headway by appealing to 'facts' about the meaning of  $\alpha$ . And so on, until we get further up the alphabet and find our best attempts at meaning isolation force us seriously to question whether we were really right about all of  $\alpha$ ,  $\beta$ ,  $\gamma$ , etc. and to re-examine our methods of abstraction and idealization, which may have led us to oversimplify in ways we now worry about. Certainly it takes some skill to keep all of the balls in the air in a stable configuration.

## 5. Pronouns and Anaphora

I have been struck, when discussing or lecturing on linguistic pragmatism, how often people assume it implies, or at least strongly suggests, a 'pragmatic' theory of pronouns along the lines of those recommended by Chomsky (1976), Lasnik (1976) or Bach (1987, 1994).<sup>56</sup> According to such theories, the third-person pronouns in the following sentences are interpreted in a uniform manner:

- (1) He's up early
- (2) Mary loves him
- (3) John's wife loves him
- (4) John loves his wife
- (5) Mary thinks John loves his wife

---

<sup>55</sup> If there is a worry about pragmatism, it surely resides here and not in minutiae like relativization (implicit binding).

<sup>56</sup> This label is employed by Evans (1980) and Bach (1987, 1994).



- (6) John thinks his wife loves him.

They all function as referring expressions, it is claimed. In each case the speaker is free to choose any male he likes. Even John. No special *co-referential* or otherwise *anaphoric* use of ‘he’ or ‘him’ or ‘his’ is exemplified by any of these sentences. Co-reference is a ‘pragmatic’ matter. John is just one of the legitimate referents for an indexical use of, say, ‘his’ in (4). And in such cases ‘his’ and ‘John’ are co-referential. No rule of grammar is involved, and to that extent the phenomenon of co-reference is not illuminated by talk of *anaphora*, construed as a grammatical notion.

At first blush, the pragmatic theory of pronouns is economical, relying, as it does, on a single reading of, say, ‘his’. It is, however, forced to posit a second, *bound* reading of pronouns, to be employed when they are hooked up to quantified expressions, and heroic efforts would be needed to reduce binding to indexicality. More importantly, the pragmatic theory just makes the wrong empirical predictions in very many cases and is, in any case, unilluminating of the way pronouns are used.<sup>57</sup> At the heart of the right theory, I maintain, is the thesis that every pronoun *seemingly* anaphoric on a name  $\alpha$  (or any other referring expression) is either (i) *bound* by  $\alpha$  (in the usual semantic sense) or else (ii) an occurrence of an indexical that is merely co-referential with  $\alpha$ . It will take some time to work up to the sort of general theory I think will be satisfactory, but it will help to see where I am going if I state three of the main theses immediately.

1. *Binding*. The pronouns in (4)-(6) *do* admit of readings upon which they are anaphoric on ‘John’; and on those readings they are *not* co-referential with ‘John’ but *bound* by it in precisely the same way they are bound by ‘every man’ on the analogous readings of (4’)-(6’):<sup>58</sup>

- (4’) every man loves his wife  
 (5’) Mary thinks every man loves his wife  
 (6’) every man thinks his wife loves him.

More generally, if a pronoun  $\beta$  is genuinely anaphoric on an expression  $\alpha$ , that is because  $\alpha$  binds  $\beta$ . It is important to appreciate that although the analysis of (4) according to which ‘his’ is bound by ‘John’ is equivalent an analysis according to

---

<sup>57</sup> On this matter, I find myself largely in agreement with points made by Geach (1962), Evans (1977, 1980), Partee (1975), Sag (1976), Soames (1989, 1990, 1994), Williams (1977).

<sup>58</sup> The proposal is essentially Geach’s (1962).

which ‘his’ is *de jure* co-referential with ‘John’, this is merely an artefact of the example. The corresponding analyses are not equivalent for sentences (5) and (6).<sup>59</sup>

2. *Co-reference*. From the fact that anaphora always involves binding, it follows that co-reference plays no rôle whatsoever in a theory of pronominal anaphora. Anaphora involves only binding. Reference and co-reference are pragmatic notions (as the ‘pragmatic’ theory says). Speakers use names, demonstratives, descriptions, and pronouns to refer to people, places, and things. Furthermore, sometimes they use such expressions, particularly pronouns, to refer to things they have already referred to. In such cases we get co-reference. We can call this ‘pragmatic anaphora’ or ‘discourse anaphora’ or ‘unbound anaphora’ if we like; but let us not confuse the notion with linguistic, i.e. bound anaphora.<sup>60</sup>

3. *Determiners*. Following Postal (1966) and others, third person pronouns are definite determiners. We must be careful, then, to distinguish the determiner [<sub>DPhe</sub>] from the full determiner phrase [<sub>DP</sub>[<sub>DPhe</sub>][<sub>NPe</sub>]], where the NP *e* is aphonic. The entire DP is in effect a definite description, hence a quantificational expression.

In the next few sections, I am going to abstract from speakers’ intentions in order that we may focus on constraints on the interpretation of pronouns imposed by grammar. When the time is right, the picture of pronouns that develops will be set in a broader pragmatist picture.

## 6. Pronouns in Generative Grammar

One weakness in early work on pronouns undertaken by philosophers was inherited from the dominant ‘descriptive’ picture of reference that would be overturned in the early 1970s.<sup>61</sup> Another stemmed from failures to appreciate ways

<sup>59</sup> Partee (1975), Soames (1990, 1994). I am not here concerned with the *de dicto* readings of readings of (5) and (6), the truth of which do not require anyone to actually have a wife.

<sup>60</sup> If this is right, the labels ‘anaphora’ and ‘anaphoric pronoun’ are ill-chosen, and we would be better off just talking about *binding* and *bound pronouns*. In *the New Shorter Oxford English Dictionary* we find the following under ‘anaphora’: ‘*Rhet.* The repetition of the same word or phrase in several successive clauses . . . *Ling.* The use of an expression which refers to or stands for an earlier word or group of words.’ In *Webster’s* we find this: ‘*Rhet.* repetition of a word or words at the beginning of two or more successive verses, clauses, or sentences . . . *Gram.* the use of a word as a regular grammatical substitute for a preceding word or group of words.’ (Webster’s goes on to give us an example: ‘as the use of *it* and *do* in *I know it and he does too.*’) As far as pronouns are concerned, perhaps ‘anaphora’ would be more properly applied where we have what Geach (1972) calls *pronouns of laziness*. Be that as it may, the word is used widely in linguistics and philosophy and there is no turning back the clock.

<sup>61</sup> Kripke (1980) and Kaplan (1989).

in which grammatical structure constrains scope possibilities and thereby bound variable readings of pronouns. By contrast, the principal weakness in early work undertaken by linguists stemmed primarily from failures to appreciate logical, semantic, and epistemic distinctions that seemed self-evident to philosophers weaned on a common fare of logical puzzles and truth definitions for formal languages. The situation in both disciplines improved dramatically in the 1970s and 1980s as a result of fertile interaction;<sup>62</sup> and a better understanding emerged of the ways in which the interpretation of pronouns reflects general organizing principles of grammar.<sup>63</sup> Work from that period is still at the centre of much current research, although grammatical theory underwent one of its periodic ‘perspective’ changes in the 1990s, the implications of which are still being examined in connection with what is known as the Binding Theory.<sup>64</sup> A potted history of how pronouns have been regarded by generative linguists will help frame the issues we need to address.

One of the central ideas in generative linguistics in the 1960s was a distinction between a sentence’s *surface* structure and its *deep* structure.<sup>65</sup> Deep structures were generated by context-free phrase structure rules such as

$$\begin{aligned} S &\rightarrow NP + VP \\ VP &\rightarrow V + NP. \end{aligned}$$

*Transformational* rules would then map deep structures into surface structures by processes that might delete, add, re-order or substitute constituents.<sup>66</sup> For example, a surface structure of roughly the form of (1) might be derived by a *passivization* transformation from something of roughly the form of (1’):

- (1) Mary was kissed by John  
 (1’) John kissed Mary.

The surface structure (2) might be derived from the deep structure (2’)

- (2) John wishes to leave  
 (2’) John<sub>1</sub> wishes John<sub>1</sub> to leave.

---

<sup>62</sup> Chomsky (1981), Cooper (1979), Evans (1977, 1980, 1982), Harman (1972), Higginbotham (1980, 1983*a*), Heim (1982/1988), Kamp (1981), Karttunen (1976), Keenan (1971), Lakoff (1972) Lasnik (1976), Lewis (1975), McCawley (1976), May (1977), Partee (1972), Reinhart (1976, 1978), Sag (1976), Wasow (1972), Williams (1977).

<sup>63</sup> Chomsky (1981), Higginbotham (1980).

<sup>64</sup> Chomsky (1995, 2000), Reuland (2001*a,b*).

<sup>65</sup> Chomsky (1964, 1965).

<sup>66</sup> Strictly, deletions and additions suffice to produce re-orderings and substitutions, but it was easier to work with more complex transformations.

by a *deletion* transformation on the basis of an *identity* in the latter, the presence of two distinct occurrences of the same noun phrase, ‘John’, marked in some way in grammar as co-referential, perhaps using integers as indices.<sup>67</sup> That identity of form was required seemed clear from the fact that the deletion transformation could not apply to, say, ‘John wants Mary to leave’; that identity of interpretation was required seemed evident from the fact that (2) cannot be used to say that John Lennon wants John Wayne to leave.<sup>68</sup>

It seemed natural within this framework to explore the idea that anaphoric pronouns were the superficial manifestations of fuller noun phrases, derived by *pronominalization* and *reflexivization* transformations.<sup>69</sup> The surface structures (3) and (4) (on their anaphoric readings), for example, were taken by some linguists to be derived transformationally from the deep structures (3′) and (4′), respectively (on the basis of noun phrase identity in those deep structures):

- (3) John<sub>1</sub> thinks he<sub>1</sub> is the best person for the job  
 (3′) John<sub>1</sub> thinks John<sub>1</sub> is the best person for the job

---

<sup>67</sup> Chomsky (1965), Postal (1966). The non-synonymy of (2) and (2′), and related pairs, was not appreciated by linguists at the time. Consider,

- (i) Mary thinks John wishes to leave.

For an utterance of this to be true Mary must believe that John satisfies *x wants x to leave*. It is not enough that Mary believe he satisfies *x wants John to leave*. Furthermore, any adequate account of the original sentence (2) must reflect the fact that an utterance of it cannot be true unless (roughly) John conceives of the experiencer of his wish and the agent of the wished-for event (or at least the agent of an event of the wished-for type) as identical. (Stephen Schiffer has suggested to me that this is too strong and over-intellectualizes what is involved. I think he may be right, but there is an itch here that needs scratching.) This is not so for every utterance of (2′). On a closely related note, consider the non-synonymy of (ii) and (ii′), discussed in detail by Fodor (1975: 133-41):

- (ii) Only Churchill remembers giving the speech about blood, sweat, toil, and tears.  
 (ii′) Only Churchill remembers Churchill giving that speech about blood, sweat, toil, and tears.

Contemporary syntactic theory maintains that in both (i) and (ii) the embedded clause, infinitival in (i), gerundive in (ii), has as its subject an anaphoric pronoun PRO. The contrast between (ii) and (ii′), as well as that between (2) and (2′) in the text, shows that the interpretive properties of PRO must differ from those of names. Moreover, the contrast between (ii) and (ii′) shows that the interpretation of PRO must differ from those of reflexive pronouns:

- (ii′′) Only Churchill remembers himself giving the speech about blood, sweat, toil, and tears.

Any adequate account of (ii) must reflect the fact that an utterance of it cannot be true unless (roughly) Churchill conceives of the experiencer of the memory and the agent of the remembered event as identical. (Again, Schiffer’s worries loom large.) For extensive discussions of these and related issues, see Castañeda (1966, 1967, 1968), Cherchia (1990), Fodor (1975), Higginbotham (1990), Lewis (1979), Partee (1975), Perry (1979, 2000), Salmon (1986, 1992), and Soames (1990, 1994). The combined force of these works establishes two incontrovertible results. First, not all pronouns are the product of transformations; second, anaphora cannot be reduced to *de jure* co-reference. These results will loom large in what is to come, but the reader is asked to bracket them for the purposes of following the potted history in this section.

<sup>68</sup> Chomsky (1965), Postal (1966).

<sup>69</sup> Langacker (1969), Lees and Klima (1963), Postal (1966), Ross (1967).

- (4) John<sub>1</sub> loves his<sub>1</sub> wife  
 (4') John<sub>1</sub> loves John<sub>1</sub>'s wife

It was soon clear, however, that pronominalization could not provide an explanation of *all* anaphoric pronouns.<sup>70</sup> First, anaphora on quantified phrases was a problem. Transformational rules were supposed to be meaning-preserving.<sup>71</sup> Thus (5) (on its anaphoric reading) could not be derived from (5'), as they clearly differ in meaning:

- (5) [every man]<sub>1</sub> loves his wife.  
 (5') [every man]<sub>1</sub> loves [every man]<sub>1</sub>'s wife.

Second, infinite regresses were discovered in so-called 'crossing co-reference' or 'Bach-Peters' structures like (6):<sup>72</sup>

- (6) [the pilot who shot at it<sub>2</sub>]<sub>1</sub> hit [the MiG that was chasing him<sub>1</sub>]<sub>2</sub>.

(6) contains two anaphoric pronouns, 'it' and 'he', each of which appears to be anaphoric on an expression containing the other. If these pronouns are the products of pronominalizing 'the MiG that was chasing him' and 'the pilot who shot at it' respectively, then (6) derives from (6'):

- (6') [the pilot who shot at [the MiG that was chasing him<sub>1</sub>]<sub>2</sub>]<sub>1</sub> hit  
 [the MiG that was chasing [the pilot who shot at it<sub>2</sub>]<sub>1</sub>]<sub>2</sub>.

But (6') contains occurrences of 'him' and 'it' that need to be derived. If they are derived by pronominalization too, obviously we have an infinite regress on our hands.

In the light of the considerations just adduced and others, by the mid-1970s it was generally accepted that at least some anaphoric pronouns were 'base-generated', i.e. present at deep structure rather than derived by a pronominalization transformation. Around the same time, the idea that there might be general constraints (or conditions) restricting the nature of transformations was being explored.<sup>73</sup> Soon enough, pronominalization was abandoned in favour of the idea all pronouns were base-generated.<sup>74</sup> And soon enough, the original idea that

---

<sup>70</sup> Bach (1970), Fodor (1975), Helke (1971), Jacobson (1977), Jackendoff (1972), Karttunen (1969, 1976), Lasnik (1976), McCawley (1976), Partee (1975), Wasow (1972). That pronouns anaphoric on quantified phrases could not be interpreted as repetitions of their antecedents was already recognized by Geach (1962).

<sup>71</sup> Katz and Postal (1964), Chomsky (1964, 1965).

<sup>72</sup> Bach (1970).

<sup>73</sup> Ross (1967), Chomsky (1973).

<sup>74</sup> Jackendoff (1972), Wasow (1972).

anaphoric relations were marked in syntax was itself called into question: the reading of a sentence upon which a (non-reflexive) pronoun  $\beta$  is seemingly anaphoric on some expression  $\alpha$  is nothing more, it was suggested, than a reading upon which nothing in the grammar *precludes*  $\alpha$  and  $\beta$  from being co-referential.<sup>75</sup> The interesting task was seen to be that of specifying precisely the syntactic conditions that precluded co-reference in examples such as those marked with an asterisk below:

- (7) John<sub>1</sub> thought he<sub>1</sub> would win
- (8) \* he<sub>1</sub> thought John<sub>1</sub> would win
- (9) his<sub>1</sub> mother thought John<sub>1</sub> would win
- (10) although he<sub>1</sub> was tired, John<sub>1</sub> could not sleep
- (10) \* Mary told him<sub>1</sub> about John<sub>1</sub>
- (11) \* Mary told John<sub>1</sub> about him<sub>1</sub>

However, by 1980 the wisdom of giving up the idea that anaphoric relations were marked in syntax was being called into question by at least philosophers working on syntax and semantics.<sup>76</sup> To talk of conditions on *co-reference*, it was pointed out, was to talk of conditions on a *symmetric* relation. But the anaphoric relations at issue are inherently *asymmetric*. Talk of *de facto* co-reference, and even talk of intentional co-reference, needed to be replaced, it was argued, by talk of *referential dependence*, a species of *de jure* co-reference. Co-reference is *not* actually precluded in the examples marked with asterisks above; what is precluded is *referential dependence*.

It is my belief that all talk of *de jure* co-reference and referential dependence is ultimately misplaced in a theory of anaphora. In accordance with where we began in our discussion of pronouns, let us say that what is precluded in the relevant examples above is *binding*, leaving it open for now whether binding is a primitive notion, as many people believe, or one that can be analysed in terms of some form of *de jure* co-reference as, for example, Evans (1977, 1980) holds. Thus I shall talk of both quantifiers and names binding pronouns, deferring until later precisely how a unified theory is to be elaborated.

---

<sup>75</sup> Lasnik (1976), Chomsky (1976), Reinhart (1976).

<sup>76</sup> Evans (1977, 1980), Higginbotham (1980).

## 7. 'Binding' and 'Scope'

One of the main aims of generative grammar, according to Chomsky, is to render explicit and systematic what the native speaker-hearer implicitly knows about the structure of his or her language simply in virtue of *being* a native speaker-hearer, and thereby to shed light on the nature of the human language faculty. We acquire language because we have a *language faculty*, a component of our shared biological endowment that unfolds in accordance with a preset programme under the triggering and shaping effect of linguistic experience. Whilst the final state may differ across individuals—particularly in so far as their exposure is to different speaker-hearers in different communities—the shaping effect of experience is quite limited. And it is this limited variation alone that leads us to talk informally about people being speakers of 'different languages'.<sup>77</sup> Although the study of conditions (or constraints) on rules of grammar had begun in the 1960s, it was not until the early 1980s that the emphasis in generative linguistics shifted dramatically from rules for generating (and interpreting) particular linguistic structures to constraints on possible structures and their interpretations.<sup>78</sup> Diverse linguistic phenomena, seemingly governed by intricate rules that differed from language to language, were now viewed as consequences of the interaction of general principles of the human language faculty, principles that were meant to be invariant across typologically distinct languages, superficial differences between particular languages reflecting only the setting of different values to each of a batch of structural parameters as part of the process of language acquisition, and the peripheral effects of relatively unimportant, learned idiosyncrasies. The interpretation of pronouns was prominent in this work because of the importance within the emerging theory of a sub-theory that concerned itself with the *binding* of one expression by another.<sup>79</sup>

The conception of binding involved in the Binding Theory is often said to be *syntactic*. But it would be quite incorrect to say it is *purely* syntactic (except on a very broad and ultimately unhelpful use of 'syntactic') because it has a clear interpretive dimension. (The nature of this interpretation, via the notion of abstraction, will be examined later.) The usual Binding Theory definition of binding might be put as follows:

$\alpha$  binds  $\beta$  iff (i)  $\alpha$  and  $\beta$  are *co-indexed*, and

---

<sup>77</sup> For a highly readable and comprehensive account of Chomsky's work, its implications, and the evolution of the core grammatical ideas, see Smith (1999).

<sup>78</sup> Chomsky (1977, 1981).

<sup>79</sup> Chomsky (1981, 1986).

(ii)  $\beta$  is within the *scope* of  $\alpha$ .

Of course, much turns on what is meant by ‘scope’ here. If the word merely labels a precise relation that holds between points of a tree, then it is easy to see why this definition of binding is called syntactic: whether or not  $\alpha$  binds  $\beta$  in a sentence  $S(\alpha, \beta)$  is something that can be determined by simply consulting  $S(\alpha, \beta)$ ’s syntactic structure, i.e. its LF, assuming minimalism. Nonetheless, we must not overlook the fact that our hypothesis about  $S(\alpha, \beta)$ ’s LF was formed partly on the basis of judgments of an *interpretive* nature; for a sentence’s LF, as Chomsky has consistently maintained, incorporates ‘whatever features of sentence structure (i) enter into the semantic interpretation of sentences and (ii) are strictly determined by properties of sentence grammar.’

Let us take scope first. In the first-order predicate calculus, a variable  $\beta$  may be bound by a quantifier  $\alpha$  (in a formula  $X$ ) only if  $\beta$  resides in *the smallest sentence (open or closed) (in  $X$ ) that contains  $\alpha$* . Following Russell, this is precisely what we mean when we say that a variable may be bound by a quantifier only if it lies within the quantifier’s *scope*. The only non-atomic expressions in the calculus are whole sentences (open or closed) and the only expressions whose scopes we care about are the sentence operators ( $\forall x$ ), ( $\exists x$ ),  $\sim$ ,  $\bullet$ ,  $\vee$ ,  $\supset$ , and  $\equiv$ . So we say, with seeming generality, that an expression  $\beta$  is within the scope of an expression  $\alpha$  iff  $\beta$  resides in the smallest *sentence* containing  $\alpha$ . In short:

For any *sentence operator*  $\alpha$ ,  $\alpha$ ’s scope = the smallest *sentence* properly containing  $\alpha$ .<sup>80</sup>

The temptation to simplify and fully generalize is hard to resist:

For any *expression*  $\alpha$ ,  $\alpha$ ’s scope = the smallest *constituent* properly containing  $\alpha$ .<sup>81</sup>

<sup>80</sup> In the terminology of syntactic theory,  $\beta$ ’s being within the scope of  $\alpha$  amounts to  $\alpha$ ’s *commanding*  $\beta$  in Langacker’s (1966) sense:  $\alpha$  commands  $\beta$  iff the minimal S node dominating  $\alpha$  also dominates  $\beta$ .

<sup>81</sup> In the terminology of syntactic theory again,  $\beta$ ’s being within the scope of  $\alpha$  now amounts to  $\alpha$ ’s *c-commanding*  $\beta$  in Reinhart’s (1976, 1978) sense:  $\alpha$  *c-commands*  $\beta$  iff the first branching node (of whatever category) dominating  $\alpha$  also dominates  $\beta$ . ( $\alpha$  and  $\beta$  are assumed to be non-overlapping). The route to c-command in linguistics was tortuous. In effect, *c-command* superseded Lasnik’s (1976) less general notion of *kommand*:  $\alpha$  kommands  $\beta$  iff the minimal cyclical node (S or NP) dominating  $\alpha$  also dominates  $\beta$ . (NP in Lasnik’s definition corresponds to what we call DP today). *Kommand* itself was meant to supersede Langacker’s (1966) still less general notion of *command*:  $\alpha$  commands  $\beta$  iff the minimal S node dominating  $\alpha$  also dominates  $\beta$ . The route through command and kommand to c-command is somewhat surprising given that the last is the simplest and most general. As Evans (1977) notes, what amounts to a c-command account of scope was in fact employed back in 1964 by Klima, who was investigating the scope of negation and *wh*-phrases. Klima offers a conditional not a biconditional: ‘A constituent . . . is “in construction with” another if the former is dominated by (that is, occurs somewhere lower down the branch of) the first branching node . . . that



Not only does this generalization appear right for natural language, minimalist assumptions appear to explain *why*: it is *beautifully idiotic*. Why is it that (1) below is perfectly acceptable, whilst (2) is not?

- (1) John told Paul's mother a lot about himself
- (2) \* John's mother told Paul's mother a lot about himself.

That is, what prevents 'himself' being bound by 'John' in (2)? The lay answer is that 'John' is not the *subject* of (2), 'John's mother' is.<sup>82</sup> However, while there are, in fact, languages in which a reflexive must be bound by a subject expression (e.g. German and Icelandic), English appears not to be such a language, witness (3):

- (3) John<sup>1</sup> told Paul<sup>2</sup> a lot about himself<sub>1/2</sub>.

Here, 'himself' can be bound by 'John' or 'Paul' (I indicate binding possibilities by placing a numerical superscript on the binder and a corresponding subscript on the bound).<sup>83</sup> Although the lay answer to our question appears to be technically incorrect, it is *right in spirit*. The *subject* of a sentence S is the DP that combines with a VP to form S (the DP 'immediately dominated' by the S node). This simple fact gives us *everything we need* to understand scope. We can characterize scope 'inclusively' (in the manner familiar to philosophers and logicians), or 'exclusively' (in a manner more familiar to linguists):<sup>84</sup>

If  $\alpha$  and  $\beta$  merge to form  $\{\alpha \beta\}$ , then  $\{\alpha \beta\} = \text{the scope of } \alpha = \text{the scope of } \beta$ .  
("My sister and I form my scope.")

If  $\alpha$  and  $\beta$  merge to form  $\{\alpha \beta\}$ , then  $\alpha = \text{the scope of } \beta$  and  $\beta = \text{the scope of } \alpha$ .  
("My sister is my scope").

I say this is 'beautifully idiotic' because *scope is something you get when you merge expressions*. And from a combinatorial or computational point of view, this makes the notion virtually trivial, something that arises as a matter of virtual conceptual necessity once we accept the possibility (as we must) of expressions

---

dominates the latter' (1964: 297). Evans puts Klima's notion to use in straightforwardly characterizing when a pronoun is within the scope of (and hence bindable by) a quantifier expression. See also Wasow (1972), Wexler, Culicover and Hamburger (1974), and Culicover (1976).

<sup>82</sup> Since mothers are female, there is no acceptable reading of (2). By contrast, brothers are male, so there is an acceptable reading of (i), despite the fact that 'himself' still cannot be bound by 'John':

- (i) John's brother told Paul's mother a lot about himself.

<sup>83</sup> In Icelandic, the two readings of (3) are captured using two different pronouns. See below.

<sup>84</sup> The *inclusive-exclusive* terminology was suggested to me by James McCawley. The word 'merge' is used in an intuitive sense here, but it is meant to foreshadow talk in the next section of its use within Chomsky's (1995) minimalist framework.

merging to form larger expressions.<sup>85</sup> It is a good methodological assumption to view ourselves as computational idiots, as beings who require the representations over which computations are defined to be as simple as possible. It is hard to imagine anything more in line with the idiot assumption than scope.<sup>86</sup>

It is now easy to see what is going on in (1)-(3). In (1), ‘himself’ lies within the scope of ‘John’ but not within the scope of ‘Paul’; in (2) it lies within the scope of neither; and in (3) it lies within the scopes of both.

We turn now to *indexing*. To say that  $\alpha$  and  $\beta$  are *co-indexed* is to say something of syntactic import that may or may not have interpretive consequences, depending upon how indexing is elaborated. Let us suppose each DP is assigned some index, which we might indicate with a subscript.<sup>87</sup> What prevents co-indexing being of ‘merely syntactic’ interest (if this even makes sense) is its *interpretation*. The whole point of indices would disappear if co-indexing were not meant to indicate something of *interpretive* significance. When ‘himself’ takes the same index as ‘John’ in (3) we think of the two expressions as linked for purposes of interpretation. One option is to say that the linking involves *co-reference*, the reflexive being *referentially dependent* upon the name. But let us continue with the idea that the linking actually involves *binding* in the sense familiar from quantification theory, with the aim of providing a uniform treatment of (3) and (3’)

(3’) [every bishop]<sup>1</sup> told [some prince]<sup>2</sup> a lot about himself<sub>1/2</sub>.

When ‘himself’ takes the same index as ‘John’ in (3), it is being used to say something that is true iff the following condition is true of (is satisfied by) some person John:<sup>88</sup>

(4)  $x$  told Paul a lot about  $x$ .

When ‘himself’ takes the same index as ‘Paul’, it is being use to say something that is true iff the following condition is true of (is satisfied by) some person Paul:

---

<sup>85</sup> So why ‘branching nodes’ in Reinhart’s original definition of c-command? Because she and others wanted to allow for the possibility of *non*-branching nodes, i.e. nodes that immediately dominate a single node. If we wanted to allow for the possibility of non-branching nodes, we could make the obvious adjustments. If a good case can be made that all non-terminal nodes are branching—perhaps even binary-branching, as seems very plausible to some linguists—then reference to non-uniqueness and branching nodes could be dropped. On my understanding of the minimalist program, the idea of a non-branching node makes no sense.

<sup>86</sup> On this matter, I take myself to be largely in agreement with Reinhart (1983, 2000) and Reuland (2001*b*).

<sup>87</sup> As Chomsky (1995) points out, such a use of indices is not obviously consistent with his minimalist assumptions.

<sup>88</sup> To say this does not preclude the possibility that in more interesting cases these two characterizations come apart.

(4') *John told x a lot about x.*

If we did not have distinct interpretations in mind, we could never have even reached the point of bringing indices into the picture. And to this extent it would be misleading to say that binding and co-indexing are *purely syntactic* notions. It is quite clear that the principal phenomenon we are investigating is an *interpretive* one with a *syntactic* dimension. (*Any discussion* of the syntactic conditions governing anaphora is up to its neck in matters of *interpretation*, for tautologically that is *precisely* what the facts that are being accounted for involve.) And to this extent talk of the syntactic conception of binding being a *purely syntactic* notion, and talk of facts about binding being *purely syntactic* can be taken with a large grain of salt.

## 8. The Binding Theory

Chomsky's Binding Theory comprises three 'locality' principles meant to characterize certain syntactic constraints on interpretation, particularly in connection with pronouns. It aims to capture the fact that (A) a reflexive pronoun (e.g. 'himself') must not be 'too far' from its binder; that (B) a non-reflexive (e.g. 'him') must not be 'too close'; and that (C) a name (broadly construed) must not have a binder at all.

In the 1980s and 1990s, great effort was expended on defining 'too far' and 'too close' in connection with binding. To a first, rough approximation—revisions will appear later—the first of the three principles that make up the Binding Theory, Principle A, requires that the antecedent  $\alpha$  of a reflexive pronoun  $\beta$  lie in  $\beta$ 's own clause (i.e. within the smallest clause containing  $\beta$ ):

*Principle A:* a reflexive is to be interpreted as bound by an expression in its own clause.

Consider (1) and (2) (asterisks before subscripts signal impossible bindings):

- (1) [<sub>S</sub> John<sup>1</sup> says that [<sub>S</sub> [no barber]<sup>2</sup> shaves himself<sup>\*1/2</sup>]]  
 (2) [<sub>S</sub> John<sup>1</sup> says that [<sub>S</sub> Paul<sup>1</sup> shaves himself<sup>\*1/2</sup>]].

The reflexive 'himself' in utterances of these sentences is bound by the subject of the embedded clause—'no barber' in (1), and 'Paul' in (2)—not by the subject of the larger clause ('John', in both examples). It is common in linguistics to say that pronouns satisfying Principle A are "locally bound."<sup>89</sup> (It is also common to avoid

<sup>89</sup> It is not difficult to find apparent violations of Principle A once we turn to other languages. In the Icelandic (i), for example, either *Jón* or *Páll* may bind the reflexive *sig*:

calling them “pronouns” and call them “anaphors”, but I eschew the former stricture here.)

By contrast, the non-reflexive ‘him’ in utterances of (1’) and (2’) below *cannot* be bound by the subject of the embedded clause, but *may* be bound by the subject of the larger clause:

- (1’) [S John<sup>1</sup> says that [S [no barber]<sup>2</sup> shaves him<sub>1/\*2</sub>]]  
 (2’) [S John<sup>1</sup> says that [S Paul<sup>2</sup> shaves him<sub>1/\*2</sub>]].

These data are in accordance with the second principle of the Binding Theory:

*Principle B*: a non-reflexive is not to be interpreted as bound by an expression in its own clause.<sup>90</sup>

The third and final principle concerns names (broadly speaking):

*Principle C*: a name is not to be interpreted as bound.

Thus, an utterance of (3) may not be understood in such a way that ‘Tully’ is bound by either ‘Cicero’ or ‘a barber’:

- (i) *Jón<sup>1</sup> segir að [Páll<sup>2</sup> raki sig<sub>1/2</sub>]*  
 John says that Paul shaves-SUBJ self-ACC  
 (‘John<sup>1</sup> says that Paul<sup>2</sup> shaves him<sub>1</sub>/himself<sub>2</sub>’).

The interesting question here is whether this and related cases (in Chinese for example) undermine the core concepts of the Binding Theory, merely push in the direction of explicable revisions, or indicate a lexical division between two different types of reflexive, only one of which satisfies Principle A. (The general issue is discussed in a way that philosophers will find highly congenial by Baker (2001).) The matter is complicated by various factors. First, the Icelandic violation is not possible when the clause containing *sig* is *indicative*—it must be subjunctive, as in (i), or infinitival as in (ii) (where PRO is the aponic subject of the embedded infinitival clause):

- (ii) [*Jón<sup>1</sup> taldi Páll<sup>2</sup> á [PRO<sub>2</sub> að raka sig<sub>1/2</sub>]*]  
 John persuaded Paul-ACC shave-INF self-ACC  
 (‘John<sup>1</sup> persuaded Paul<sup>2</sup> [PRO<sub>2</sub> to shave him<sub>1</sub>/himself<sub>2</sub>]’).

(I have skated over certain matters here that are discussed later.) Second, Icelandic reflexives are subject-oriented, by which is meant that they may be bound only by subject expressions. (There are interesting exceptions, however.) Third, only the monomorphemic reflexive *sig* and its case variants (*sér* (DAT), *sín* (GEN)) engage in principle A violations of this sort. The corresponding complex reflexives (*sjálfan sig*, *sjálfum sér*, *sjálfs sín*) are locally bound. Fourth, some verbs (e.g. *raka sig* ‘shave oneself’) appear to be intrinsically reflexive and to resist complex reflexives. This collection of characteristics is not uncommon for long-distance reflexives across languages and certainly makes one wonder whether there are not simply two types of reflexive pronoun in natural languages; see Baker (2001), Cole et al. (2001), Pica (1987, 1991). Icelandic pronouns have been investigated in detail by Bresnan (2001), Maling (1984, 1986), Reuland (2001a), Reuland and Sigurjónsdóttir (1997), Sells (1987), Sigurðsson (1990), Sigurjónsdóttir (1992), Þráinsson (1976, 1990, 1991), and others.

<sup>90</sup> There are certain predicates that may require (or at least strongly favour) a particular non-reflexive bound pronoun as an argument. Higginbotham (1980) gives an example similar to (i):

- (i) John brought a date with him.

The origins of Principle B are in Chomsky’s (1976) rule of non-co-reference.

(3) [<sub>S</sub> Cicero<sup>1</sup> says that [<sub>S</sub> [a barber]<sup>2</sup> shaves Tully<sub>\*1/\*2</sub>]].

To say this is not to say that ‘Cicero’ and ‘Tully’ may not be *de facto* co-referential.<sup>91</sup> It is just to say that ‘Tully’ cannot be bound by ‘Cicero’. On the assumption that the concept of binding involves interpretive as well as syntactic notions, it is unclear why the issue of binding a name should come up in the first place, and this might cast doubt on the need for Principle C.<sup>92</sup>

Despite all manner of counterexamples, the Binding Theory still holds centre stage in discussions of pronouns and anaphora—although syntactic theory itself has hardly stood still—a fixed point from which to explore. Providing a clear and precise specification of the (seemingly) complementary distribution of reflexives and non-reflexives with respect to their binding possibilities, one that holds across a multitude of languages would be a phenomenal accomplishment, of course. But even if a version of the Binding Theory were to emerge that everyone found acceptable, it might still fall well short of where Chomsky wants us to be, for what he wants (and surely we should agree with him here) is not just a *description* of the facts, but an *explanation* of *why* the Binding Theory holds. The theory would certainly not be ‘necessary’ in any sense of this word usually employed by philosophers; but this does not mean it may not be ‘conceptually’ or ‘empirically’ necessary in Chomsky’s (1995, 2002) sense: if a grammar is the optimal solution to the problem of relating form and meaning for a system as expressively rich as the ones we do, as a matter of empirical fact, possess, it may turn out the interpretive difference between ‘himself’ and ‘him’ (or, rather more plausibly, between two more general classes of expressions) is pretty much unavoidable. As Chomsky might put it, it is a matter of ‘virtual conceptual necessity.’<sup>93</sup> Since there is no condition on variables in first-order logic corresponding to the locality difference described by the Binding Theory, its existence in natural language is something that needs an empirical explanation, as it *appears* natural language is making things more complicated than it might.<sup>94</sup>

One terminological point can be made immediately. The word ‘pronoun’ tends to be used in a special technical way by Chomsky and many other linguists so as to

---

<sup>91</sup> Evans (1980).

<sup>92</sup> The origins of Principle C are in Lasnik’s (1976) rule of non-co-reference, and as such it appears to be a hangover from a time when linguistics ran together the symmetric notion of *de facto* co-reference and the asymmetric notion of referential dependence. See Evans (1977, 1980).

<sup>93</sup> See Reuland (2001*b*).

<sup>94</sup> My own suspicion about the source of the locality bifurcation enshrined in Principles A and B has to do with the concept of predication.

exclude what we have been calling ‘reflexive pronouns’.<sup>95</sup> Along with reciprocals (‘each other’), reflexives fall under the label ‘anaphor’ for Chomsky. Principle A is meant to govern the interpretation of ‘anaphors’, so a more general formulation would contain ‘anaphor’ where we put ‘reflexive pronoun’ earlier. The non-reflexive ‘him’, by contrast is a ‘pronoun’ (or ‘pronominal’) for Chomsky. Since Principle B of the Binding Theory is meant to govern the interpretation of ‘pronouns’, a proper formulation would contain just ‘pronoun’ where we put ‘non-reflexive pronoun’. This usage will not be adopted here for two reasons. First, we want to ensure continuity with the philosophy literature (and, indeed, a good portion of the linguistics literature, where there is frequent talk of ‘unbound anaphora’), so we shall continue to use ‘pronoun’ in the broad way, distinguishing reflexive and non-reflexive forms. Second, it is perfectly consistent with Chomsky’s thinking that his anaphor-pronoun distinction—an anaphor being an expression that falls under Principle A, and a pronoun one that falls under Principle B—is at best a taxonomic artefact, one that may be useful in our standard bootstrapping practice, but one that finds no place in a final theory.<sup>96</sup>

## 9. ‘Pragmatic’ and ‘Ambiguous’

Before looking more closely at binding, I want to embed the picture that is developing within the broader pragmatist picture of interpretation with which I began. Evans (1982: 309) rightly stresses there is ‘no infallible linguistic guide’ to determining when understanding what someone is saying on a given occasion requires grasping a thought about a particular individual, for ‘most expressions which are conventionally apt for such a use also have other uses’. In this connection, he immediately cites pronouns, on the grounds that they may be used to make independent reference to individuals (rather like demonstratives) or as the natural language counterparts of bound variables. When philosophers and linguists discuss expressions with more than one use, the adjectives ‘ambiguous’ and ‘pragmatic’ are soon in the air. Where pronouns are concerned, we might easily fall into using these words in a number of ways that, although related, are not identical, and we must take not to cross wires.<sup>97</sup> One very common use of ‘pragmatic’ is pointed out by Evans (1980). If *A* utters,

---

<sup>95</sup> Chomsky (1981, 1986, 1995).

<sup>96</sup> Reuland (2001*a,b*).

<sup>97</sup> Every now and then one comes across a self-serving edict on the ‘correct’ or ‘best’ uses of ‘semantic’ and ‘pragmatic’. But these words have been used in so many related ways in the philosophy of language and linguistics that any detailed attempt to explore, separate, and legislate is

- (1) *He's up early*

referring to a man who is walking by, or

- (2) *I'm glad he's left*

referring to a man who has just left, the references of (*A's* utterances of) the third person pronouns are determined by 'what may loosely be called "pragmatic" factors' (1980/1985: 216) and not by 'reference to *the rules of the language*' (1980/1985: 230). What Evans means here is that neither the meaning of the word 'he' nor any rule of sentence grammar (nor the combination) determines who the pronouns in (1) and (2) are being used to refer to in the situations sketched:

All that the conventions governing the referring expression 'he' insist upon, in any given context, is that the object referred to should be male (1982: 312). There is no linguistic rule which determines that a 'he' or a 'that man' refers to *x* rather than *y* in the vicinity, or that it refers to someone who has just left rather than someone who has been recently mentioned.' (1980/1985: 230-1).

In the language of the early part of this chapter, who *A* is referring to by 'he' in the imagined utterances of (1) or (2) is determined by *A's* referential intentions. And these intentions are constrained by the (known) conventions governing the referring use of 'he' as well as all the other sorts of things discussed earlier. As soon as we introduce anaphoric pronouns matters become more complicated, but not unmanageably so. In accordance with Principle A of the Binding Theory, the reflexive 'himself' *must be* bound. In accordance with Principle B, the non-reflexives 'he', 'him', and 'his' *can be* bound. As stated, the principles put reflexive and non-reflexives in complementary distribution as far as binding possibilities are concerned.

Consider:

- (3) John<sup>1</sup> told Paul's<sup>2</sup> wife a lot about himself<sub>1/\*2</sub>

It would be a mistake to say the speaker had a choice here about which expression he intends to be understood as binding 'himself'. Among other things, what the conventions governing the use of 'himself' seem to insist upon, in any given context, is that it be bound by a masculine DP that stands to it in a configuration satisfying Principle A. (To be sure, this is not how an ordinary speaker would put things, but I take that to be irrelevant.) Principle A, once formulated to linguists' satisfaction, precludes 'Paul' from playing that rôle.

---

pointless and perverse. The best one can hope for is consistency in one's own usage—if not across an entire body of work at least within some useful portion—with occasional pointers that serve to differentiate one's usage from others with which it might be confused.

It is not difficult, however, to construct from (3) a case in which Principle A does not force a unique result. Compare (3) with (4), where ‘himself’ may be bound by either ‘John’ or ‘Paul’:

- (4) John<sup>1</sup> told Paul<sup>2</sup> a lot about himself<sub>1/2</sub>.

And this fact might occasion another use of the word ‘pragmatic’. If it is false that the reflexive’s binder is uniquely fixed by ‘reference to *the rules of the language*’, it might be suggested, surely we need *some* help from ‘what may loosely be called “pragmatic” factors.’<sup>98</sup> But we must be careful here. Earlier I stressed the asymmetric, indeed dovetailed, epistemic positions of the speaker and the hearer, and I separated, and then reconnected, the rôles of the rules of the language and the rôles of the speaker’s intentions in theoretical talk about *what determines what a speaker is saying* and about *what determines who or what a speaker is referring to*. If we are not careful to bear all this in mind we will find ourselves with a classic case of two wrongs engendering the illusion of no wrong whatsoever, for the contrast between (3) and (4) brings out rather nicely a pair of quite distinct confusions which, like many other pairs, may manage to avoid detection for a good while by effectively drawing attention away from one another, screening one another off, as it were.

To cement ideas, I want to examine a simple example in some detail. Early one evening, I uttered the following sentence to someone who had just arrived in Reykjavik and called me on the telephone:

- (5) He is going to collect you at seven o’clock, then come over here for me.

Let us assume, for simplicity, that there is no interesting lexical or structural ambiguity we need to examine in connection with (5). The sentence itself does not reveal fully what I intended to say because (for starters) it doesn’t tell you who I meant by the pronoun ‘he’; nor, contrary to the way some philosophers seem to write on this topic, does it tell you who I meant by ‘you’, ‘here’, or even ‘me’.

---

<sup>98</sup> The ambiguity in (4) is not mirrored in all languages. In Icelandic, for example, only an expression in subject position may bind a reflexive:

- (i) *Jón<sup>1</sup> sagði Páli<sup>2</sup> frá (sjálfum) sér<sub>1/\*2</sub>*  
 John told Paul-DAT about (EMPH) self-DAT  
 (‘John<sup>1</sup> told Paul<sup>2</sup> about himself<sub>1/\*2</sub>’)

(The simplex reflexive in Icelandic inflects for case, but not for number or gender: *síg* (ACC), *sér* (DAT), *sín* (GEN); there is no nominative form. In (i) it is assigned dative case by the preposition *frá*.) To produce an Icelandic sentence equivalent to the reading of (4) upon which ‘his’ is bound by ‘Paul’, one would have to use a non-reflexive pronoun—one might even use the equivalent of ‘him Paul’. The matter is rather more complicated than I am making out, however, for reasons that will emerge.



Evans's (1982) discussion is helpful here (see esp. 9.2, pp. 309-20). On his account, the conventions governing a referring expression  $\alpha$ , as uttered on a particular occasion, associate with  $\alpha$  a property which an individual must satisfy if it is to be the referent of 'a fully conventional use' of  $\alpha$ , on that occasion. For example, the utterance of 'I' or 'me', on a given occasion, is associated with the property of *being the person making the utterance*; and the utterance of 'you', on a given occasion, with the property of *being the person being addressed in that utterance*. (1982: 311). The conventions governing 'he' and 'this' and 'that' are much less discriminating, Evans notes. And, this fact, he rightly says, 'precludes any simple functional definition of the notion of 'a referring expression in an utterance' as the expression which indicates which object the speaker is talking about.' (1982: 312, n. 9). In many cases, he notes, the *predicate* used may do more for the hearer 'to narrow down the range of interpretations of the referring expression than does the referring expression itself' (*ibid.*).

Four notions are frequently invoked at this point in discussions: (i) *context of utterance*, (ii) *salience* (iii) *Gricean or other pragmatic principles*, and (iv) *speaker's intentions*. And unfortunately they are often introduced as components of a single entity, as if they somehow conspire to 'bridge the gap' between the meaning of a sentence and what a speaker says on a given occasion by uttering that sentence. As we saw earlier, category mistakes will be our reward if we fail to separate the perspectives of speaker and hearer, but if the earlier discussion is respected, we shall have no trouble.

What is it *I know* about my utterance, beyond knowing the meaning of the sentence I used (or at least the meanings of its words and their syntactic arrangement), that makes it the case that *I know* what I intended to say? Answer: *I know* to whom I intended to refer by 'he' (Mike), 'you' (Don), and 'me' (Stephen), and I know what I intended to refer to by 'here' (my apartment). I intended to say to Don that Mike was going to collect him at seven o'clock (that same evening), then come over to my apartment to collect me.

But *how* do I know to whom or to what I intended to refer by 'he', 'you', 'me' and 'here'? Answer: I have privileged access to my referential intentions. They were, in fact, perfectly *respectable* referential intentions in the circumstances because I could reasonably expect Don to recognize that I intended to refer to Mike, Don, myself, and my apartment, respectively. (The situation would not have been dramatically different if I had used 'Mike' instead of 'he'.)

I was directly aware of my communicative intentions. I know who or what I intended to refer to by 'he', 'you', 'me' and 'here', and what I meant quite

generally. Scanning the environment is not going to provide me with any information that can help me identify what I meant, though of course my awareness of the immediate environment or context *constrains* the referential and more generally communicative intentions I can reasonably have given my choice of words. The formation of genuine intentions is severely constrained by belief, and a speaker's referential intentions are no exception, being constrained by knowledge of word meaning and syntax, tacit beliefs about who was speaking, who was being addressed, who was salient in the conversational context, and much more besides. But nothing like a 'dynamic salience metric' or a formal 'context' or 'index' 'fixed' or 'determined' who and what I intended to refer to in the example just given; for, tautologically, *that* was determined by *my referential intentions*, and it is these intentions it is the task of the hearer to identify.<sup>99</sup>

To understand the nature of the constraints on the formation of referential intentions, let us switch quickly to the perspective of the hearer. Don heard me utter (5). What did *he need to know* about my utterance, beyond knowing the meaning of the sentence I uttered (or at least the meanings of its words and their syntactic arrangement), in order to identify what I intended to say? Answer: he needed to establish whom I intended to refer to by 'he' (Mike), 'you' (Don), and 'me' (Stephen), and where I was referring to by 'here' (my apartment). And since Don had no direct access to my referential intentions, he had to establish these things *in some other way*.

Let us begin with 'he'. We no more need to say that 'he' is ambiguous in order to make sense of 'he' being used with the intention to refer to different individuals on different occasions than we need to say that 'I' is ambiguous. When I uttered 'he' as part of my utterance of (5), I intended to refer to Mike. And this referential intention was implicated in a *communicative* intention I had: to convey to Don that Mike was going to collect him at seven o'clock (that same evening), then come over to my apartment for me. In fact, I intended Don to understand me as *saying* that Mike was going to collect him at seven o'clock then come over to my apartment for me. And I could not reasonably be credited with such an intention if I thought it unlikely that Don would understand me as intending to say that. Why was I prepared to put so much faith in a use of the word 'he'? Because (i) I know  $\#$  it is part of what one knows when one knows the meaning of the word 'he' that it is to be used to refer to a male, (ii) I know Don speaks English, (iii) I tacitly take Don to be operating in accordance with shared general pragmatic principles, and (iv) Don had just asked me a question about Mike—something like 'I can't get

---

<sup>99</sup> As pointed out earlier (see n. 30) in certain circumstances one may seek to disguise one's intended referent.

through to Mike. What's the plan for tonight? Is he picking us up?'—making Mike, whom we both know to be male, an object of conversation (salient). So I used the pronoun 'he' intending to refer to Mike and expecting Don to *recognize* this. Exactly *how* this happens is something a cognitive theory of utterance interpretation (such as Relevance Theory) must tell us. And although we do not yet have a complete story about the cognitive processes involved, we do seem to have a plausible philosophical account of *what* happens, in the sense of what the net result is. That's what our *judgments* tell us.

Neither context nor salience nor pragmatic principles nor word meaning nor syntax 'fixed' or 'determined' who I intended to refer to by 'he', for tautologically that was determined by *my referential intention* in using it. Which is not to say that context, salience, pragmatic principles, word meaning and syntax played no rôle. They played two dovetailed, mutually sustaining roles. They were things that *Don used* (in different ways to be sure) in identifying my referential intention. But they were also things that *I used* because my beliefs (tacit or otherwise) about context, salience, pragmatic principles, word meaning and syntax are things that bear on the formation of referential intentions.

We must distinguish, then, (a) *what determines who I intended to refer to by 'he'* and (b) *what Don uses in identifying who I intended to refer to by it*. There is only misery in store for those who would confuse these notions. Don used the meaning of the word 'he', its syntactic position, and information about the context of utterance to identify my referential intention.<sup>100</sup> And my referential intention was constrained by the position and meaning of the word 'he', which I tacitly assumed Don also knew, by my tacit beliefs about who was salient and about the pragmatic principles we both employ, and so in sum by my estimation of Don's ability to recognize that, in this particular context, I would be intending to refer to Mike.<sup>101</sup>

---

<sup>100</sup> If I had said 'He told Mike to pick you up . . .' the position of 'he' would have steered Don away from interpreting me as referring to Mike by 'he' (assuming only one person named 'Mike' is relevant to the exchange). Indeed, given the question Don asked, he would have struggled to come up with a sensible interpretation.

<sup>101</sup> Whenever the notion of intention is introduced, there are philosophers who immediately bring up error. What if the person the speaker intends to refer to by 'he' is female? What if the intention turns out to involve no-one at all? What if the person *A* intends to refer to by 'Smith' is not Smith but Jones? What if there is no object that *A* intends to refer to by 'that' because he has been fooled by a reflection or is hallucinating? What if *A* meant to utter the word 'brother' rather than 'father' but misfired? It is here that disputes may arise, among ordinary speakers and philosophers alike. And it would seem that as far as a philosophical theory of interpretation is concerned, we will have to see where theoretical developments take us. My own strong inclination is to say that such problem cases illustrate conflicts *among intentions* not a problem with the intentional story *per se*. Certainly we should not construct our theories primarily with a view to producing clear results in cases where things go wrong and where we ourselves are not clear. As Chomsky, Davidson, and Grice have

So our respective perspectives on the exchange are not independent. In producing my utterance, I relied on what I took to be Don's capacity to identify what I intended to convey, an intention partly shaped by my estimation of that capacity. And Don assumed that I was so relying. And, possibly, so on. In short, the ways in which Don and I were operating were dovetailed and mutually sustaining, and this is quite general, cases of breakdown only reinforcing the point.

If we think about this a little more, we see it has little to do with pronouns or indexical expressions *per se*. Suppose I utter a sentence *X* that no-one has ever uttered before, a sentence whose words are all individually unambiguous and free of contextual variability (perhaps there is no such a sentence, in which case my point is already made). And suppose that these features of *X* ensure that wherever and whenever anyone utters *X* he or she will always be saying that *p* and so always be taken to be saying that *p* by any rational person who hears the utterance correctly, though I, the speaker, have never reflected on this fact. There is still something I intend to be taken as *saying* when I utter *X*, and I reasonably expect that I will be so taken. My hearer does not have direct access to this communicative intention and is, as ever, in the position of having to identify it on the basis of my utterance and information obtained from context. I expect him to succeed because I assume he knows the meanings of the words in *X* and their syntactic arrangement and can hear me clearly and exploit contextual information, and so on. It just so happens that, by hypothesis, in this particular case no *additional* information that may be extracted from context will have any impact whatsoever on the hearer's task. If the case is implausible, so much better for my point.

The situation with respect to my uses of 'you', 'here', and 'me' in my utterance of (5) differs only in the nature of the constraints the meanings of these words impose on viable referential intentions. My knowledge of the meanings of 'you', 'here', and 'me' involves knowing that they constrain my referential intentions to be directed towards (roughly) my audience, my location, and myself, respectively, giving me considerably less leeway than with the third person 'he'. When I uttered 'you', 'here', and 'me' in uttering (5), I exploited the fact that Don knew the meanings of these words and would, in consequence, take my referential intentions to be directed towards him, my location, and myself, respectively.

Again, it is important to see that context does not *determine* who I am referring to by 'you', 'here', and 'me'. My intentions determine these things, but the

---

stressed, in their own ways, let us not lose sight of the many cases in which things go right, the cases that motivated our inquiries in the first place. As far as the present chapter is concerned, nothing of consequence to my main point turns on cases of error.

intentions I may have are constrained by my beliefs about the context of utterance, about what I know about the meanings of ‘you’, ‘here’, and ‘me’, about what I assume Don knows about them, and so on. The context of utterance is important not because it determines who I am referring to but because it serves up a body of information used by Don to identify who I am referring to, a body of information I, the speaker, tacitly assumed he would draw upon.

Let us now return to (4):

(4) John<sup>1</sup> told Paul<sup>2</sup> a lot about himself<sub>1/2</sub>.

Imagine the simplest case. The speaker, *A*, is fully aware that the rules of the language require him to use ‘himself’ in his utterance of (4) in such a way that it is bound by either ‘John’ or ‘Paul’. And imagine that *A* has operated in accordance with these rules and that his hearer, *B*, knows he has (let’s say *A* even *tells* him as much). *B* is still not home yet. It is here that ‘context’, ‘intentions’, ‘salience’, and ‘pragmatic principles’ and so on are usually brought into the picture, and it is important to get the picture right. Let us suppose that *A* uttered ‘himself’ intending it to be understood as bound by ‘John’, the latter used to refer to, say, John Lennon. *A* intended to say something about John Lennon, viz. that he satisfied the condition

(6) *x* told Paul a lot about *x*

(or that the predicate ( $\lambda x(x \text{ told Paul a lot about } x)$ ) applied to him). In order to do this, *A* selected (4). Why such confidence in the reflexive pronoun? Because *A* knows, and assumes *B* knows, that the rules of language require ‘himself’ to be bound by ‘John’ or ‘Paul’ when uttering this particular sentence. And given what *A* and *B* have just been talking about—the way John Lennon has been telling people about himself today, let us suppose—*A* is quite justified in supposing *B* will interpret ‘himself’ as bound by ‘John’. In summary, *A* has a *binding* intention.<sup>102</sup> For simplicity, let us say that a binding intention is a special form of referential intention (without committing ourselves to the position that ultimately binding must be (or is best) explained in terms of co-reference (a position which seems to me false).

The question of what determines the interpretation of ‘himself’ in an utterance of (4) is one that resists certain simplistic answers found in the literature—

---

<sup>102</sup> If the position, to be adopted later, that the binding of a pronoun  $\beta$  by an expression  $\alpha$  amounts to the merging of  $\alpha$  with a  $\lambda$ -predicate whose operator binds  $\beta$ , formally spelling out what goes on when ‘Paul’ is meant to be binding ‘himself’ in (4) is a little tricky, perhaps indicating why there is a tendency to subject-orientation of reflexives in many languages. This matter is discussed later in connection with the existence of locality constraints.

‘syntax’, ‘context’, ‘salience’, ‘pragmatic principles’. If there is a short answer at all it is ‘the speaker’s intentions’. But the formation of particular intentions is severely constrained because of *A*’s tacit understanding of such things as syntax and pragmatic principles, and his beliefs (tacit or otherwise) about who or what is salient. Given *A*’s knowledge of the way ordinary middle sized objects behave, he cannot sensibly intend to jump over the Empire State Building; and given his knowledge of the rules of the language, he cannot sensibly intend to use ‘himself’ in (4) in such a way that it is bound by neither ‘John’ nor ‘Paul’. In the envisaged scenario, *A* assumes that the famous pragmatic factors will lead *B* to plump for ‘John’.

From the perspective of the hearer, simplistic answers will not do either. *B* seeks to interpret *A*’s remark and will fail unless he identifies which name *A* intends to be understood as binding ‘himself’. He has no direct access to *A*’s binding intentions, but there are lots of things he can get his hands on, as well as an important fact entailed by the rules of the language: on the assumption that *A* has not misspoken, ‘himself’ must be bound by either ‘John’ or ‘Paul’. So grammar narrows down the options to two, and the famous pragmatic factors enable *B* to narrow these down to one.<sup>103</sup>

At the beginning of this chapter I said that a major problem that has plagued the study of pronouns is that all kinds of superficial generalizations and distinctions can be made, that it is difficult to establish which are of theoretical significance and which are mere distractions, and just as difficult to establish which of various competing classifications and unifications are going to pay theoretical dividends and which are ultimately going to lead to dead-ends. This is particularly true when we examine disputes about ambiguity. Here there is no quick way to tell when we are onto something of deep theoretical significance, or even when there is a genuine dispute in any given case: all we can do is attempt to construct a general theory of pronouns (as part of an overall theory of interpretation) and see where we find the most useful unifications and bifurcations, not just within a language but also across languages—at least if we are serious about explaining linguistic structures in terms of general principles of the language faculty and thereby shedding light on the logical problem of language acquisition.

In order not to prejudge too many issues, let us continue using the phonologists’ old slash notation when talking about the superficial form of a string of words. Consider,

---

<sup>103</sup> I am not articulating a psychological theory of what is taking place in the heads of *A* and *B*, of course, but rather giving a philosophical reconstruction of what a psychological theory must explain, in much the same way as Grice does in his philosophical psychology.

(7) /every man loves his wife/.

(7) contains what is often called the possessive or genitive pronoun /his/, and thereby the possessive or genitive description /his wife/.<sup>104</sup> It is frequently said that (7) is ambiguous, because /his/ admits of two quite distinct uses, one on which it is bound by ‘every man’, another on which it is used to make independent, indexical reference to some salient male.<sup>105</sup> If this description of the situation is accurate, then one of the tasks facing a hearer, *B*, when presented with an utterance of (7), is to identify whether the speaker, *A*, is using /his/ in the indexical or bound way. And even if the distinction between these two uses of /his/ is encoded *grammatically* there is still nothing in *the superficial form* (7) to indicate which use *A* intends; so *B* must use *non-linguistic* information to establish how *A* is using /his/. And in *this* respect, interpreting an utterance of (1) is a *pragmatic* matter, on one perfectly useful and sensible use of the adjective ‘pragmatic’. Nothing changes if we take these two uses of /his/ to correspond, strictly speaking, to the use of *two distinct words* pronounced /his/.<sup>106</sup>

That said, *B*’s knowledge of word meaning and syntax *do* tell him something important about any utterance of (7), though *B* may not have reflected consciously on this: that *A* may be using /his/ in one of two distinct ways, that two different *types* of interpretation may be associated with utterances of (1). It is surely part of what *B* knows by virtue of knowing English that *A* may legitimately intend /his/ to be read as making independent reference to some particular individual or as bound by ‘every man’. This is *not* a pragmatic matter. *B* knows this fact about (7) without appealing to any non-linguistic information. It is a *linguistic* fact—a syntactico-semantic fact—that these two uses of (7) are possible.

<sup>104</sup> As far as English is concerned, it appears not to matter whether we draw a distinction between the possessive and the genitive. This is not true for all languages, however. See the discussion of Icelandic below.

<sup>105</sup> I should stress that there exist theories according to which there is no ambiguity or dual use of the sort I have just described, at least in cases in which the pronoun has a singular antecedent such as (i):

(i) John loves his mother.

Lasnik (1976), for example, and following him Bach (1987) argue that the occurrence of ‘his’ in (i) refers to some salient male. On this account, the two uses of ‘his’ we have posited correspond to the choice of different salient male, in one case someone already mentioned in the sentence, in the other someone not so-mentioned. This theory is thoroughly undermined by points made by Evans (1980) and Soames (1990).

<sup>106</sup> Analogy: interpreting utterances of /John is at the bank/ and /visiting professors can be a nuisance/. (But see below.) The possessive /his/ brings up an orthogonal interpretive problem: the nature of the intended relation. There are many possibilities for, say, /his horse/, which is why I used the nominal /wife/ in the example above. (Actually, all sorts of possibilities are possible for /his wife/, but you probably didn’t realize that when you first looked at (7).) Why didn’t I use /he/ or /him/ to circumvent distracting side issues about possession? Because using /his/ enables me to make an important point I cannot make using the others. See the discussion of Icelandic below.

## 10. Minimalism, LF, and Binding

In the 1990s, syntactic theory evolved considerably in the light of *minimalist assumptions*, the net effect of which was to restrict the posits of grammatical theory—whether categories, processes, constraints, or levels of representation—to those that are conceptually necessary or empirically unavoidable.<sup>107</sup> A motif that runs through this work is an argument from ‘virtual conceptual necessity’: complexity and stipulation are to be avoided as, all else being equal, language will employ only those devices needed to link sound and meaning. On the assumption that there is a component of the mind/brain dedicated to language, the human language faculty, one consequence of the minimalist outlook is that all properties of sentences relevant to sound and meaning—and this includes *binding* properties—should be derivable from quite general considerations about the way the language faculty must engage with two other cognitive systems, one dealing with the articulation of sounds and their perception (henceforth *the sound system*), the other trading in intentional/conceptual representations (henceforth *the intentional system*). A particular language can be seen as an instantiation of the language faculty (with certain options specified), something that can provide ‘instructions’ to be interpreted by the sound system, on the one hand, and the intentional system, on the other. More specifically, a language is a computational system that generates pairs  $\langle \pi, \lambda \rangle$  of representations, where  $\pi$  is a PF (or ‘Phonetic Form’) to be read by the sound system, and  $\lambda$  an LF (or ‘Logical Form’) to be read by the intentional system. (Where early generative grammar distinguished the Deep Structure and Surface Structure of a sentence, and later versions distinguished its D-Structure, S-Structure, PF, and LF, minimalism allows for just PF and LF.)

For convenience we can identify a sentence with a pair  $\langle \pi, \lambda \rangle$ . For Chomsky, the basic idea behind the concept of LF representations has remained robust since its inception: An LF incorporates, ‘whatever features of sentence structure (1) enter into the semantic interpretation of sentences and (2) are strictly determined by properties of sentence grammar’ (Chomsky, 1976b: 305). The only difference today is what is meant by ‘strictly determined by properties of sentence grammar’. With the emergence of the minimalist outlook, this phrase may be usefully understood as ‘strictly determined by the exigencies of connecting sound and meaning.’

---

<sup>107</sup> Chomsky (1995, 2002).



As Chomsky has stressed for a quarter of a century, LFs are not full-fledged *intentional* representations, not as rich in content as the those involved in beliefs, intentions or expectations, not objects with truth conditions (or the analogues thereof), inferential rôles and so on. LFs are simply the grammar's contribution to the generation of such representations by the intentional system—which receives inputs from various cognitive faculties. To get the flavour of this idea it is helpful to think of the intentional system as trading in representations something like the sentences of the 'language of thought' in Fodor's (1975, 1983) sense, a modality-neutral symbolic system of representation in which thought takes place and into whose sentences utterances of natural language sentences must be mapped if understanding is to take place. LFs exhaust the grammar's contributions to this system; an LF is not an interpreted object with an intentional content: it is simply a syntactic representation, determined by sentence grammar, of those features of grammatical structure that enter into the interpretation of utterances of that sentence (for example, relations of scope and binding). Not only do LFs fail, for example, to specify references for referentially independent occurrences of pronouns, they fall short of being full-blown intentional representations in all sorts of other ways.<sup>108</sup>

Reference to the theory of phrase structure is virtually eliminated in the minimalist framework, there being no phrase structure rules in the traditional sense

---

<sup>108</sup> See Chomsky (1975, 1976, 1986, 1995, 2002), Sperber and Wilson (1986, 1995), and Carston (2002). In a Davidsonian spirit, Higginbotham and May (1981), Higginbotham (1980, 1983a, 1983b), Larson and Segal (1995), Ludlow (1989), Neale (1994) and others have treated LFs as objects which (relative to assignments of values to referential elements, some of which may be of an indexical nature) have recursively specifiable truth conditions, an idea Chomsky rejects. There is, I think, *something* of real importance in the alternative conception of LF that needs to be super-imposed upon the official Chomskyan conception to give it bite, but it is difficult to make this precise. In trying to effect the superimposition, my own earlier discussions of LF evince deeply worrying ambiguities. In 'Events and LF' (1988) and *Descriptions* (1990) the discussion is very Chomskyan: I treat LFs as no more than syntactic objects encoding those aspects of syntax relevant to interpretation, and I am careful to distinguish LFs themselves from the formulae in a system of restricted quantification I use to represent the truth conditions of utterances of sentences, formulae that could depart significantly in structure from the LFs of the sentences uttered (for example, in the discussions of perceptual reports, incomplete descriptions, and descriptive pronouns). Various people tried to convince me I needed to embrace, or at least move closer to the truth-evaluable conception of LFs, but the interpretation of *utterances* containing incomplete descriptions and other underspecified DPs, as well as the interpretation of D-type pronouns held me back as it seemed, and still seems, preposterous to treat LFs as containing all sorts of unrecoverable *predicative* material not present in surface syntax. At the same time, it has always seemed to me that if the Chomskyan LF of a sentence *X* is meant to lay bare the contribution made to the process of *interpretation* by *X*'s syntax, it must provide genuine constraints on what someone uttering *X* can be *saying*, something for which the alternative truth-conditional conception is tailor-made. In 'Logical Form and LF' (1994), I attempted a reconciliation of sorts—originally with Larson—by abstracting as much as possible from *pragmatically* determined aspects of the truth-conditions of utterances. The attempt was ultimately unsuccessful, I think, because I strayed so far from the Chomskyan conception of LF.

and, strictly speaking, no constituent structure to a PF representation. Again we can borrow the old /slash/ notation from phonology when talking about PFs, but with standard orthography rather than phonological symbols inside the slashes, thus individuating coarsely in terms of phonological properties:

- (1) /John said he was at the bank/  
 (2) /every man loves his mother/.

We can still allow ourselves the convenience of using a phrase structure tree or a labeled bracketing to explicate superficial structure, so to speak, even if what we write down is strictly a hybrid of PF and LF. In contemporary syntactic theory, expressions such as ‘some man’, ‘a man’, ‘the man’, ‘every man’, and so on are usually called DPs (determiner phrases) to reflect the idea that they are projected from the Ds (determiners) ‘some’, ‘a’, ‘the’, ‘every’ and so on (rather than from the N (noun) ‘man’ as earlier theory suggested).<sup>109</sup> On this usage, which we shall follow, the label NP is reserved for the nominal expression, simple or complex, with which the determiner merges to form a DP, as in the following tree and labeled bracketing:<sup>110</sup>

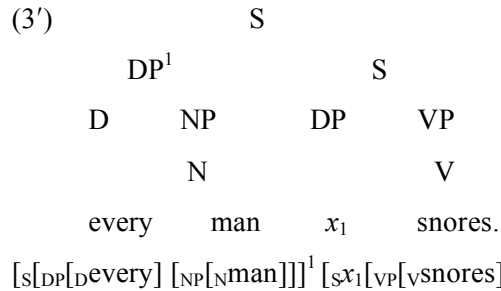
- (3)
- |  |       |         |
|--|-------|---------|
|  | S     |         |
|  | DP    | VP      |
|  | D     | NP      |
|  |       | V       |
|  | N     |         |
|  | every | man     |
|  |       | snores. |
- [<sub>DP</sub>[<sub>D</sub>every] [<sub>NP</sub>[<sub>N</sub>man]]] [<sub>VP</sub>[<sub>V</sub>snores]]]

The LF corresponding to the PF (3) will look something like (3'):<sup>111</sup>

<sup>109</sup> For discussion, see Abney (1987), Cardinaletti (1994), Cardinaletti and Starke (1999), Chomsky (1995), Elbourne (2001, 2002), Szabolcsi (1994). In the elementary exposition of DPs that follows, I simplify dramatically as certain on-going debates in linguistics about the details are not crucial to the philosophical issues of concern here.

<sup>110</sup> Prior to the work of Abney (1987), what are called DPs today were usually called NPs to reflect the idea that they were projections of Ns.

<sup>111</sup> Higginbotham (1980, 1983a), Higginbotham and May (1981), May (1977, 1985).



In (3') the quantifier expression 'every man' has been extracted—indeed forced by general principles of morphosyntax—from its original position (discernible in the *mélange* (3)) and has merged with the original S node to form another. From this 'new' position it binds the variable  $x$  that has been left as a sort of 'trace' in its 'original' position, that position being within the scope of its new position. The numerical *subscript* on  $x$  indicates that it is to be interpreted as bound by the quantifier expression 'every man', which bears the same index as *superscript*.<sup>112</sup> This talk of variables and binding amounts to a description of an important part of the *interpretive* information carried by (3), precisely the sort of thing Chomsky ascribes to LFs. Variables are expressions *interpreted* in a certain way.

One question that will have to be addressed is whether quantifier expressions are the *only* DPs that are raised at LF in this way or whether the phenomenon is fully general, involving names, pronouns, and possessives for example (DPs whose structures we will look at shortly).<sup>113</sup>

- (4) [<sub>S</sub>[<sub>DP</sub>John]<sup>1</sup>[<sub>S</sub>[<sub>DP</sub>Ann]<sup>2</sup>[<sub>S</sub> $x_1$ [<sub>VP</sub>loves  $x_2$ ]]]]]  
 (5) [<sub>S</sub>[<sub>DP</sub>he]<sup>1</sup>[<sub>S</sub>[<sub>DP</sub>Ann]<sup>2</sup>[<sub>S</sub> $x_1$ [<sub>VP</sub>loves  $x_2$ ]]]]]  
 (6) [<sub>S</sub>[<sub>DP</sub>John]<sup>1</sup>[[<sub>DP</sub>his<sub>1</sub> wife]<sup>2</sup>[<sub>S</sub> $x_1$ [<sub>VP</sub>loves  $x_2$ ]]]]].

It is at least clear from (6) that if the possessive description 'his wife' is raised, then 'John' will have to be raised too if it is to bind 'his'.

When the Binding Theory was first formulated, a sentence was not factored into just its PF and LF. D-Structure and S-Structure were *bona fide* levels of grammatical and so it was possible to debate where the Binding Theory had to apply. Some advocated LF, others S-Structure; some suggested it must apply at both levels, others at at least one; and examples were even found that suggested D-Structure might be the appropriate level. Within a minimalist framework such

<sup>112</sup> Using only subscripts to co-index would obscure the fact that binding is an asymmetric relation. I am deliberately simplifying here. For example, I have ignored the fact that the superscripted index on the DP 'every man' has been projected upwards from the index on D 'every'.

<sup>113</sup> If, as seems plausible, possessives are definite descriptions, and definite descriptions are quantifier expressions, then we already have our answer in one case.

debate is impossible. The Binding Theory can be only a set of constraints imposed by virtually unavoidable facts about an interpretive system, a set of principles operative at the interface of the language faculty and the system of conceptual-intentional representations, i.e. it holds at LF. (It should be noted, however, that Chomsky himself has suggested the Binding Theory disappear under minimalist assumptions. I remain unconvinced.)

## 11. Pronouns as Determiners

A proposal from 1960s linguistics that has been resuscitated to great effect recently is Postal's idea that pronouns are determiners.<sup>114</sup> Whilst this idea might seem odd at first, reflection reveals it to be rather intuitive and well-motivated.

Traditional grammars distinguish sharply between the possessive (or genitive) determiners in (1) and their absolute possessive (or genitive) counterparts in (1'):

- (1) *our/your/his/her/John's/their* car is the one on the left  
 (1') *ours/yours/his/hers/John's/theirs* is the one on the left.

In many cases it is as if the characteristically genitive 's' falls away when the determiner occurs with no overt NP.<sup>115</sup> Indeed generative linguistics suggests a single possessive determiner the superficial form of which depends upon whether it occurs with a *phonic* (phonologically non-null) or *aphonic* (phonologically null) NP complement.<sup>116</sup>

---

<sup>114</sup> Postal (1966), Stockwell et al. (1973), Abney (1987), Elbourne (2001, 2002).

<sup>115</sup> Irregularities: 'his' and 'whose' remain as they are; 'mine' and 'thine' typically become 'my' and 'thy' in modern speech.

<sup>116</sup> The postulation of expressions that are aponic despite having syntactic rôles and semantic properties is surely no more or less problematic than the postulation of expressions that are semantically empty despite having syntactic rôles and phonological properties ('it' in 'it's raining', for example). The idea of an expression that is phonetically *and* semantically empty is harder to get one's mind around, and on the interpretation of Chomsky's present framework I endorse—syntax is whatever it is that relates PF and LF—the possibility of such an expression is straightforwardly excluded. The discovery or postulation of *any* expression constitutes a contribution to syntax and its existence is justified only if it is doing something at LF or PF. Consequently, the discovery or postulation of an *aphonic* expression must be justified by its rôle at LF. To this extent, it will contribute in one way to the project of producing a theory of utterance interpretation. The point should not be exaggerated, however. Discovering or postulating occurrences of a bound, aponic expression *e* has very clear consequences for a theory of interpretation: an expression has been discovered or posited that is understood as a bound variable, effectively answering all questions about its interpretation on a given occasion. Contrast this with the discovery or postulation of an aponic, *indexical* expression, one every bit as flexible in its interpretation as the overt expressions 'this' or 'that' or 'he' (when used to make independent reference). The interpretation of an utterance of a sentence containing an occurrence of an aponic indexical is always going to be a full-fledged *pragmatic*, i.e. inferential matter, the semantics of the aponic expression itself placing only non-deterministic constraints on interpretation.

- (2) [DP<sub>D</sub>his][NP<sub>N</sub>car]] is the one on the left  
 (2') [DP<sub>D</sub>his][NP<sub>e</sub>] is the one on the left.

[NP<sub>N</sub>car]] is phonic, [NP<sub>N</sub>e]] is aphonetic. Let us distinguish, then, *absolute* ('his', 'hers') and *reliant* ('his', 'her') occurrences of determiners according as they occur with aphonetic or phonic NP complements.<sup>117</sup>

There are two ways to think about the aphonetic [NP<sub>e</sub>] serving as the complement of a possessive determiner. (i) It might be construed as the result of a process of *NP deletion* at PF.<sup>118</sup> With (3) and (3'), this seems fine:

- (3) John prefers my painting to [DP<sub>D</sub>hers] [NP<sub>e</sub>]  
 (3') John prefers my painting to [DP<sub>D</sub>her] [NP<sub>N</sub>painting]].

But if (2) and (2') derive from a single LF, we lose the recoverability of deletion.<sup>119</sup>  
 (ii) Alternatively, [NP<sub>e</sub>] might be construed as an NP that itself appears in the LF of (2') and as such is in need of pragmatic interpretation just like an occurrence of a pronoun.

By contrast with pronouns, proper names might be viewed as Ns, serving as NPs, serving as full DPs by virtue of merging (in English) with aphonetic Ds:

- (i) [DP<sub>D</sub>e][NP<sub>N</sub>John].

(In many languages (e.g. ancient and modern Greek), names appear regularly with the definite article. In certain contexts names may occur with the definite article and other determiners in English: 'That John Smith is not the John Smith I was talking about.') Perhaps we find NPs with aphonetic Ds elsewhere:

- (ii) [DP<sub>D</sub>e][NP<sub>N</sub>whales]] are [DP<sub>D</sub>e][NP<sub>N</sub>mammals]].

The following examples reveal another type of aphonetic expression, this time a whole DP:

- (iii) John promised Ann to sing  
 (iv) John asked Ann to sing.

Traditional grammars talk about the 'understood subject' of the verb 'sing' in such examples: in (iii) it is 'John', in (iv) it is 'Ann'. In syntactic theory this is captured by a difference in the status of an aphonetic element (usually called PRO) occupying the subject position of the embedded infinitival clause:

- (iii') [<sub>S</sub>John<sup>1</sup> promised Ann [<sub>S</sub>e<sub>1</sub> to sing]]  
 (iv') [<sub>S</sub>John asked Ann<sup>2</sup> [<sub>S</sub>e<sub>2</sub> to sing]].

In (iii'), the interpretation of *e* is required (by the syntax and the meaning of the verb 'promise') to proceed by way of the interpretation of the subject expression 'John' ('promise' is a *subject-control* verb). In (iv'), by contrast, the interpretation of *e* is required (by the syntax and semantics of the verb 'ask') to proceed via the interpretation of the object expression 'Ann' ('ask' is an *object-control* verb (in this construction)).

<sup>117</sup> One matter that needs to be taken up later is whether *all* occurrences of, say, /his/ have the same syntactic structure. If one ends up distinguishing bound and indexical uses of /his/, there is always the possibility that they are distinguished in virtue of syntactic structure.

<sup>118</sup> Elbourne (2001, 2002).

<sup>119</sup> It is arguable that (3)/(3') does not satisfy it either, at least if 'painting' is stressed. See Neale (2004).

Possessives are not the only determiners that have both reliant and absolute occurrences:

- (4) *many* (people) applied but *few* (people/of them) were suitably qualified
- (5) *both* (senators) spoke in favour but *neither* (senator/of them) was convincing
- (6) *some* (guests) arrived late but *some/others/most* (guests/of them) did not
- (7) *only two* (guards) fell asleep but *all three* (guards/of them) will have to be fired
- (8) *that* (piece) is geometric; *this* (piece) is Mycenaean.

The determiners ‘no’, ‘a’, ‘the’, and ‘every’ seem always to be reliant:

- (9) no/a/the/every guard fell asleep.
- (9') \*no/\*a/\*the/\*every fell asleep.

But we must not overlook the possibility, found with possessives, of morphological variation depending upon the phonicity of the complement.<sup>120</sup> For example, it is plausible that ‘none’ is the absolute form of the reliant ‘no’, and that ‘one’ is the absolute form of the reliant ‘a’ (as well as of the reliant ‘one’):

- (10) many people had dessert but *none* (of them) ordered coffee.
- (11) many people had dessert; indeed *one* (of them) had two portions.

Is there an absolute form of ‘the’? The hypothesis at hand is that there is, and that it comes in three gender variants: ‘he’, ‘she’, and ‘it’.<sup>121</sup> Third person pronouns, on this hypothesis, are actually forms of the definite article taking aphonic complements:

- (12) [DP[Dthe]][NP[Npresident]] is asleep
- (13) [DP[Dhe]][NP[e]] is asleep
- (14) Don't wake [DP[Dhim]][NP[e]].

In the first instance, the reflexive ‘himself’ might be viewed as a fusion of the determiner ‘him’ (or, perhaps more plausibly (see below), the possessive determiner ‘his’) and the nominal ‘self’:

- (16) every man loves [DP[Dhim]][NPself]],

<sup>120</sup> Elbourne (2001, 2002).

<sup>121</sup> Postal (1966), Stockwell et al. (1973), Elbourne (2001, 2002). When I was graduate student my adviser, John Perry, used to joke that I took every noun phrase to be a definite description until proven otherwise. If I had known about Postal's thesis in the 1980s, I would almost certainly have attempted to see it as the basis of a theory of descriptive anaphora of the sort sketched in Chs. 5 and 6 of *Descriptions*.

If this is adequate, and if we wish to retain the idea that reflexives and some occurrences of third person pronouns are understood as bound, it might look as though we will have to say that some occurrences of *determiners* are bound variables. But we can obtain the desired result with something interestingly weaker: a purportedly bound occurrence of a third person pronoun is just the head of a DP that *contains* a bound variable. Take a bound occurrence of ‘he’ in a sentence such as (17):

(17) [every man]<sup>1</sup> thinks that he<sub>1</sub> snores.

On the hypothesis we are considering, the occurrence of ‘he’ in (17) has the form [DP[he][NP<sub>e</sub>]]. Our task is to make sense of the idea that this DP contains (or is at least interpreted *as if* it contains) a variable bound by ‘every man’. Suppose we add to our formal language RQ a new quantificational determiner *he*. And suppose, mirroring the pronominal hypothesis we are considering, that *he* is a special form of *the* (forget about gender for a moment). Then the following trivial modification of the familiar Russellian axiom for *the* is what we want for *he*:

(HE) [*he*  $x_k$ :  $\phi$ ] $\psi$  is true of a sequence  $s$  iff  $\psi$  is true of every sequence that  $\phi$  is true of differing from  $s$  at most in the  $k$ -th position, and there is exactly one such sequence.

(If we want gender to affect truth conditions, we just insist that  $\phi$  contain *male*  $x$  as a conjunct.) The truth conditions of (17), on one of its readings, are captured by *either* of the following sentences of RQ:

(18) [*every*  $x_1$ : *man*  $x_1$ ] ( $x_1$  thinks that ([*he*  $x_2$ :  $x_2=x_1$ ] ( $x_2$  snores))))

(18') [*every*  $x_1$ : *man*  $x_1$ ] ( $x_1$  thinks that ([*he*  $x_2$ : *man*  $x_2 \cdot x_2=x_1$ ] ( $x_2$  snores)))).

If we see [NP<sub>e</sub>] as the product of NP-deletion, then we will see (17) as derived from something like (17'):

(17') [every man] realizes that [he man] snores.

And on such account, there is a natural inclination to use (18') in characterizing the logical form of (17). By contrast, if we see [NP<sub>e</sub>] as base-generated, so to speak, then (18) seems more natural. Either way, the unrestricted quantifier *he* binds the occurrence(s) of  $x_2$  in the open sentence with which it combines to form a restricted quantifier; and either way the restricted quantifier—[*he*  $x_2$ :  $x_2=x_1$ ] in (18), [*he*  $x_2$ : *man*  $x_2 \cdot x_2=x_1$ ] in (18')—binds the occurrence of  $x_2$  in  $x_2$  snores; so either way, we get a variable inside the matrix of a restricted quantifier that the other restricted quantifier [*every*  $x_1$ : *man*  $x_1$ ] binds.

We need to tie all this up with the English syntax of course; and in order to keep our options open, we need to do it for both (17) and (17'). If the structure of

the DP ‘he’ in (17) is  $[_{DP}[_{D}he\ [_{NP}e]]]$ , we get what we want if this DP is understood as equivalent to  $[he\ x_2: x_2=x_1]$  in (18), which we get if the D ‘he’—which is only *part of the DP* ‘he’—is understood as equivalent to the RQ determiner  $he\ x_2$  and if the NP  $[_{NP}e]$  is understood as equivalent to the RQ formula  $x_2=x_1$ . The morpho-syntax of English will, in fact, insist upon the complement of ‘he’ always being  $[_{NP}e]$ .<sup>122</sup> So we might schematically specify the semantics of  $[_{NP}e]$  as given by  $x_k=x_j$  (for  $j \neq k$ ) when such a device functions as the complement of the determiner ‘he’ (thus the aphonic NP in  $[_{DP}[_{D}he][[_{NP}e]]]$  has a definite semantic role here—made transparent by (18)—which would justify its existence in a minimalist framework). When we have a suitably placed co-indexed binder as in (17), we have what will amount to a *bound* occurrence of ‘he’. If there is no such binder, we have what amounts to *free* occurrence of ‘he’, used to make indexical reference to some individual.<sup>123</sup>

Talk of *bound* pronouns is still perfectly intelligible on this account. The D ‘he’ is not bound; indeed it is a binder. And the DP ‘he’ is not *wholly-bound* the way a bound variable is; rather it is just *bound-into*. The subscript on ‘he’ in (17) must now be understood as indicating that  $[_{DP}[_{D}he][[_{NP}e]]]$  is *bound-into* rather than wholly bound.

Two questions of largely notational import now arise: (i) Where should we place the subscript ‘1’ in the spelled out DP? (ii) Will we ever need a subscript ‘2’ corresponding to the ‘2’ in the RQ rendering (18)?

(18)  $[every\ x_1: man\ x_1] (x_1\ realizes\ that\ ([he\ x_2: x_2=x_1] (x_2\ snores)))$ .

In order to have something fixed, let us adopt the conventions implicit in (17'') if we are asked to set everything out in gory detail:

(17'')  $[_{DP}\ every^1\ [_{NP}[_{N}man]]]^1\ realizes\ that\ [_{DP}[_{D}he^2][[_{NP}e]_1]^2\ snores$ .

The index ‘1’ is placed on the empty NP (or at least on its node); and just as the index ‘1’ on the D ‘every’ projects to the DP ‘every man’ it heads, so the index ‘2’ on the D ‘he’ projects to the DP ‘he’ it heads.

On this account, the RQ sentences (19') and (16') might explicate the logical forms of (19) (on one of its readings) and (16), respectively:

(19) every man thinks Mary loves him

(19')  $[every\ x_1: man\ x_1] (x_1\ thinks\ that\ ([him\ x_2: x_2=x_1] (Mary\ loves\ x_2)))$

<sup>122</sup> Examples like ‘he, Chomsky, thinks that . . .’ are epenthetic hence irrelevant.

<sup>123</sup> Of course, we will need the usual constraints on indexing to prevent a DP binding into a pronoun that is meant to be free, as in (i):

(i)  $[every\ man]^1\ thinks\ he\ [free]\ likes\ him_1$ .



(16) every man loves himself

(16') [*every*  $x_1$ : *man*  $x_1$ ] ([*him*  $x_2$ :  $x_2=x_1$ ] ( $x_1$  loves  $x_2$ )).

(The axiom (HIM) for the RQ determiner *him* is just (HE) with [*he*  $x_k$ :  $\phi$ ] replaced by [*him*  $x_k$ :  $\phi$ ].)

We can probably improve upon (16') once we have dealt with the possessive 'his'. There is a long tradition in philosophy of treating possessive DPs as definite descriptions, motivated in part by the apparent equivalence of 'Smith's murderer' and 'the murderer of Smith', and so on.<sup>124</sup> So, in the first instance we might explicate the logical form of (20), on its bound reading, by treating it as superficial variant of (21), the logical form of which we can sketch using (21')

(20) every man loves his wife

(21) every man loves the wife of him

(21') [*every*  $x_1$ : *man*  $x_1$ ] ([*the*  $x_2$ :  $x_2$  wife  $x_1$ ] ( $x_1$  loves  $x_2$ )).

Now (21') is only a sketch because it contains a (mere) variable corresponding to the occurrence of the DP 'him' in (21), a view we have already gone beyond. What we really need is (21'')

(21'') [*every*  $x_1$ : *man*  $x_1$ ] ([*the*  $x_2$ : [*him*  $x_3$ :  $x_3=x_1$ ] ( $x_2$  wife  $x_3$ )] ( $x_1$  loves  $x_2$ )).

Now what of (20) itself? A desire for a sentence of RQ more closely resembling the English sentence whose semantic structure it is meant to explicate might lead one to add another determiner, *his*, to RQ:

(20') [*every*  $x_1$ : *man*  $x_1$ ] ([*his*  $x_2$ :  $x_2$  wife  $x_1$ ] ( $x_1$  loves  $x_2$ )).

The relevant axiom will be an interesting modification of the one for *he*, to be given in a moment. To bring out another feature of possessives, let us switch examples: Consider a particular utterance of (22)

(22) [every man]<sup>1</sup> groomed his<sub>1</sub> horse.

The DP 'his horse' might be understood as (e.g.) 'the horse he owned' or 'the horse he rode' or 'the horse he trained'. In RQ we might represent the DP as [*his*

<sup>124</sup> Russell (1905), Kripke (1977), Higginbotham (1983a), Neale (1990). It is arguable that some possessives are better seen as *indefinite* descriptions ('John sprained his ankle', 'my cousin is visiting'). This matter is skirted here. In some cases, the possessive must cede to the form beginning with 'the'. Thus we say 'the weight of this' not 'this's weight' (despite the acceptability of both 'the weight of this man' and 'this man's weight. This appears to lend support to the view that the demonstrative pronouns 'this' and 'that' are just occurrences of the demonstrative determiners 'this' and 'that'.

$x_k$ : horse  $x_k \bullet R(x_k, x_j)$ ] where  $R$  is *owned* or *rode* or *trained*, as the case may be.<sup>125</sup>

The following axiom would appear to yield what we want:

- (HIS) [*his*  $x_k$ :  $\phi$ ] $\psi$  is true of a sequence  $s$  iff  $\psi$  is true of every sequence that  $\phi \bullet R(x_k, x_j)$  is true of differing from  $s$  at most in the  $k$ -th position, and there is exactly one such sequence.

Of course this does not actually throw any light on the syntactic relation (if there is one) between the English sentences (20) and (21).<sup>126</sup>

Finally, the reflexive ‘himself’. We might see this as the result of combining a pronominal determiner with ‘self’, construed as a genuine nominal like ‘wife’.<sup>127</sup> If the determiner in question is just a phonetic variant of ‘his’, we can think of ‘himself’ as a possessive just like ‘his wife’ and use (16’’) rather than (16’) to explicate the logical form of (16):

- (16) every man loves himself  
 (16’) [*every*  $x_1$ : *man*  $x_1$ ] ([*him*  $x_2$ :  $x_2=x_1$ ] ( $x_1$  loves  $x_2$ ))

<sup>125</sup> No specific relation is encoded by the possessive marker *per se*. (Sperber and Wilson (1986)) At most it signals that some relation or other is to be inferred, a relation that may differ from utterance to utterance. Sometimes the relation a speaker intends can be inferred from a relation inherent in the meaning of the noun: ‘wife’, ‘mother’, ‘murderer’, ‘teacher’, and ‘mayor’, for example, bring with them relations to things or places (but see below). But many nouns forming the complements of possessive determiners do not bring with them such relations (‘donkey’, ‘man’, ‘table’); and whilst a default ownership relation may be understood in many cases, it is easy enough to find cases in which ownership is not the intended relation. I may use ‘my horse’, ‘his horse’, and ‘Smith’s horse’ to describe the horses Smith and I have staked money on in the Cheltenham Gold Cup; or to describe the horses Smith and I were riding one afternoon; and it would be stretch to say I am thereby using possessives non-literally. Furthermore, even in cases where the noun in use brings a particular relation along with it, that relation may not be the one of paramount importance to understanding what the speaker is saying. At a school sports day, Smith and I may decide to bet on the outcome of the teachers’ three-legged race or the mothers’ egg and spoon race; at a prison sports day we might bet on the outcome of the cat-burglars’ high-jump or the murderers’ 100 metres. In recounting the day’s results I may use ‘Smith’s murderer’, ‘my murderer’, ‘his mother’, ‘my teacher’, and so on quite felicitously to talk about the contestants we have bet on. (Does this mean I can use ‘my mother is not my mother’ without contradiction? Yes. Similarly someone whose birth mother is not his familial mother or legal mother. Of course we might be asked to expand on our respective remarks.)

<sup>126</sup> The LF of (20) may well be close to the LF of (21). The following is notoriously ambiguous:

- (i) this is John’s former house.

It may be used to say that (a) this house was formerly John’s or (b) that this former house is John’s (after an earthquake for example). Larson and Cho (2000) suggest this is the product of a scope ambiguity, roughly between (ii) and (iii):

- (ii) this is the former [house of John]  
 (iii) this is the [former house] of John.

Working out the details seems to require seeing the PF (i) as corresponding to two distinct LFs, roughly those corresponding to the PFs (ii) and (iii).

<sup>127</sup> On the creation of English reflexives, see Keenan (2002) and references therein. In languages other than English not just analogues of ‘self’, but names of various body parts are used in creating reflexives.

(16'') [every  $x_1$ : man  $x_1$ ] ([his  $x_2$ :  $x_2$  self  $x_1$ ] ( $x_1$  loves  $x_2$ )).

On this account, no new axiom is needed as 'himself' is just a special case of 'his NP' where the relation  $R$  is identity.<sup>128</sup> Alternatively we could introduce a new determiner *himself* whose complement is always understood as an identity:

(16''') [every  $x_1$ : man  $x_1$ ] ([*himself*  $x_2$ :  $x_1 = x_2$ ] ( $x_1$  loves  $x_2$ )).

The axiom (HIMSELF) for *himself* will involve a trivial modification of (HIM). (Of course, it does not provide an explanation of what is *special* about reflexives; that is something we still need to provide.)

The upshot of all this is that even if pronouns are determiners, as much current linguistic theory suggests, there is no barrier to making sense of the idea that some occurrences (indeed, all occurrences of reflexive pronouns, for reasons that ought to emerge) are *bound*. To say that an occurrence of, say, 'him' is bound by a quantifier  $\alpha$  is to say that the determiner 'him' is the head of a DP [<sub>DP</sub>[<sub>D</sub>him] [<sub>NP</sub> $e$ ]] whose aponic complement is understood as containing a variable bound by  $\alpha$ .

One problem people have worried about in connection with binding in syntactic theory has been silently solved. Contrast the following:

(23) [every man]<sup>1</sup> realizes that [some man]<sup>2</sup> loves himself<sub>\*1/2</sub>

(24) [every man]<sup>1</sup> realizes that he<sup>2</sup> loves himself<sub>\*1/2</sub>.

In (23), only 'some man' can bind himself, as predicted by Principle A of the Binding Theory. Similarly, only 'he' can bind 'himself' in (24). But surely 'every man' can bind 'he' in (24), just as it can in (25):

(25) [every man]<sup>1</sup> realizes that he<sub>1</sub> loves Mary.

So in (24) we need 'every man' to bind 'he', and 'he' in turn to bind 'himself'. If 'he' were a mere variable a *transitive* conception of binding might be required, the semantics of which might involve some rather fancy footwork. But on the proposal sketched here, what is needed drops out automatically, with Principles A and B respected. Qua *binder* the DP 'he' bears a superscript; qua device that is *bound-into*, its NP bears a subscript:

(23') [every man]<sup>1</sup> realizes that [<sub>DP</sub>he [<sub>NP</sub> $e$ ]<sub>i</sub>]<sup>2</sup> loves himself<sub>\*1/2</sub>.

We can usefully abbreviate this:

(23'') [every man]<sup>1</sup> realizes that he<sub>1</sub><sup>2</sup> loves himself<sub>\*1/2</sub>.

<sup>128</sup> In examples like (i) and (ii), discussed by Jackendoff (1992) and Nunberg (1977),  $R$  would be some other relation only *involving* identity:

- (i) At Madame Tussaud's, Bill Clinton took a photograph of himself.
- (ii) Yeats hated hearing himself read in an English accent.

Putting everything together, the sentence comes out as equivalent to the following, confirming that the Binding Theory has not been violated:

- (24) [every  $x_1$ : man  $x_1$ ] ( $x_1$  realizes that  
 ([he  $x_2$ :  $x_2=x_1$ ] ([his  $x_3$ :  $x_3$  self  $x_2$ ] ( $x_2$  loves  $x_3$ ))))).

A final word about the free or indexical use of pronouns before we move on. There is no obvious reason to think they differ syntactically from those that are bound (in English, at any rate). Either way, the general form is  $[_{DP}[\text{he } [_{NP}e]]]$ . But what of the interpretation of an indexical occurrence? Since the D ‘he’ is not indexical on the current proposal, the indexicality of the DP ‘he’ must lie in  $[_{NP}e]$ . The right thing to say is that whereas a bound use of ‘he’ is understood as [*he*  $x_j$ :  $x_j=x_k$ ] with  $x_k$  bound by some other DP, on an indexical use it is understood as [*he*  $x_j$ :  $x_j=x_k$ ] with  $x_k$  free to refer indexically. Thus what amounts to a unitary theory.

Now we need to unify names and quantifiers in their binding of pronouns.

## 12. Unification: Farewell to Co-reference

We have been working on the assumption that an anaphoric pronoun is one occurring within the scope of some other expression  $\alpha$  that binds it, even when  $\alpha$  is a name (or some other singular term). But, famously, many linguists and philosophers have assumed that when  $\alpha$  is a name, the pronoun functions as a device of *co-reference*. In (1), for example, it is sometimes suggested that if /John/ is the antecedent of /his/, the two expressions are co-referential:

- (1) John<sup>1</sup> loves his<sub>1</sub> wife.

While this may be a harmless assumption for many philosophical purposes, we cannot afford to be so incautious or incurious if we are in the business of providing a comprehensive theory of pronouns and anaphora. One reasonable desideratum is a uniform account of how /his/ (and derivatively /his wife/ and /loves his wife/) function in (1) and (2):

- (2) [every man]<sup>1</sup> loves his<sub>1</sub> wife.

And it is not immediately obvious how a co-reference account of /his/ in (1) and a bound variable account of /his/ in (2) can constitute a unified account.

Is it possible to reduce one of co-reference and binding to the other? According to Evans (1977, 1980) once we understand the species of co-reference we are dealing with, we can indeed explain binding in terms of it. Mere co-reference is a *symmetric* relation; but the relation we are interested in is asymmetric: on the relevant reading of (1), says Evans, /his/ inherits its reference from /John/; ‘his’ is not merely co-referential with /John/ it is *referentially dependent* upon it. In short

we are dealing with *de jure* or *rule-governed* co-reference, and it is Evans's contention that once we have (i) an account in terms of 'referential dependence' of how /his/ functions in the singular case (1), and (ii) a certain type of account of how /every man/ works in the quantified case (2), we will automatically have (iii) an account of how /his/ functions in the quantified case.

Evans's co-reference rule is straightforward: a pronoun  $\beta$  anaphoric on a singular term  $\alpha$  refers to whatever  $\alpha$  refers to.<sup>129</sup> This ensures that on the reading of (1) that interests us, the reference of 'his' is identical to the reference of 'John', and that an utterance of the whole sentence is true if and only if John loves John's wife. What of (2), the quantified counterpart of (1)? Evans opts for a 'Fregean' interpretation of quantifiers. The rule for 'every  $\phi$ ' is straightforward (if unfamiliar to those weaned on Tarski). An utterance of sentence (2) is true if and only if ' $x$  loves his wife' is true of every man; and ' $x$  loves his wife' is true of an arbitrarily selected man  $m$  if and only if, taking ' $\alpha$ ' as a name of  $m$ , the sentence ' $\alpha$  loves his wife' is true. If the language happens to contain no name of  $m$ , then we consider a minimal *extension* of the language that *does* contain a name for  $m$ ; that is, we add to our language a *new* name to use when referring to  $m$ .<sup>130</sup> Once the rule for 'every  $\phi$ ' has been applied, Evans's co-reference rule is applied, and we obtain the result that ' $\alpha$  loves his wife' is true if and only if the referent of  $\alpha$  loves the wife of the referent of  $\alpha$ , i.e. if and only if  $m$  loves  $m$ 's wife. Since  $m$  was chosen arbitrarily, this means an utterance of (2) is true if and only if for every man  $x$ ,  $x$  loves  $x$ 's wife.<sup>131</sup>

But there is a problem here, one Evans appears to recognize but whose severity he appears to underestimate. Versions of the problem have been raised by Castañeda (1966), Partee (1975), and Soames (1989, 1990, 1994). Consider the following quantified-singular pair:

- (3) Mary thinks that John<sup>1</sup> loves his<sub>1</sub> wife
- (4) Mary thinks that [every man]<sup>1</sup> loves his<sub>1</sub> wife.

<sup>129</sup> Evans himself does not actually require that  $\beta$  be within the scope of  $\alpha$ —he wants his co-reference rule to operate quite generally.

<sup>130</sup> There is no assumption here that there is some extension of the language that contains a name of *every* man, only an assumption that for every man there is some extension of the language that contains a name of him.

<sup>131</sup> See Evans (1980/1985: 227-8). As Soames (1989) points out, Evans assumes all singular terms in natural language are rigid. This is a view Evans (1982) defends. It is, of course, an empirical thesis about natural language that singular terms are rigid. The fact that one can construct formal languages that contain non-rigid singular terms or design non-rigid singular terms to be added to theoretical fragments of natural language is neither here nor there.

The bound variable treatment of the pronoun in (4) delivers (4')—with an eye to readability, henceforth I shall often use  $x$ ,  $y$ , and  $z$  in place of  $x_1$ ,  $x_2$ , and  $x_3$ , where possible:

(4') *Mary thinks that* ([every  $x$ : man  $x$ ] ( $x$  loves  $x$ 's wife)).

This captures the fact that (4) is true (on one of its readings) iff Mary thinks that every man is an own-wife's lover. Of course (4') is actually quite schematic as we are treating 'his' as a determiner. (4'') is closer to what we want:

(4'') *Mary thinks that* ([every  $x$ : man  $x$ ] ([his  $y$ :  $y$  wife  $x$ ] ( $x$  loves  $y$ ))).

The analogous reading of (3) is true iff Mary thinks that John is an own-wife's lover, but that reading is *not* captured by saying that 'John' and 'his' are co-referential.<sup>132</sup> For on such an account an utterance of (3) could be true if Mary believes that John loves the wife of some man she sees but does not realize is John.<sup>133</sup> So even if 'his' can be used as a device of *de jure* co-reference—which has not yet been established—we still appear to need to take into account its use as a device that can be bound by names. We are left with three remaining questions, then: (i) Can we provide a plausible unification in the reverse direction? (ii) If so, can we do it without treating names as quantifiers; and (iii) Can we then dispense altogether with appeals to a use of 'his' as a device of *de jure* co-reference?

(i) In order to explain what is going on in (3), we seem to need the *abstraction* introduced by the quantification in (4'). That is, we appear to need a formula in which the name 'John' binds a variable. Schematically, (3'), with more detail (3''):

(3') *Mary thinks that* ([John  $x$ ] ( $x$  loves  $x$ 's wife))

(3'') *Mary thinks that* ([John  $x$ ] ([his  $y$ :  $y$  wife  $x$ ] ( $x$  loves  $y$ ))).

This suggests we unify the treatment of pronouns bound by quantifiers and those bound by names by treating names as quantifiers.<sup>134</sup> (If one is already convinced that proper names are generalized quantifiers, or that they are disguised Russellian descriptions—then one is already persuaded. But if one has reasons for thinking names are *not* quantifiers, other possibilities will have to be explored.)

For simplicity let us go back to example (1), the logical form of which, on the present account, may be sketched using (1'):

(1') [John  $x$ ] ([his  $y$ :  $y$  wife  $x$ ] ( $x$  loves  $y$ )).

<sup>132</sup> Soames (1990, 1994), Reinhart and Reuland (1993), Heim and Kratzer (1998). The problem is replicated for other *de jure* co-reference accounts of pronouns bound by names as well as for the 'pragmatic' co-reference accounts of Lasnik (1976) and Bach (1987).

<sup>133</sup> Various assumption are made here, some of which may be non-trivial. See Soames (1990, 1994).

<sup>134</sup> On names as quantifiers, Montague (1973), Barwise and Cooper (1981).

In a Tarskian vein, the following axiom would appear to suffice for a quantifier *John*:

- (JOHN)  $[John]_k \psi$  is true of a sequence  $s$  iff  $\psi$  is true of every sequence with John in the  $k$ -th place differing from  $s$  at most in what it assigns to  $x_k$  and there is exactly one such sequence.<sup>135</sup>

We appear now to have a complete theory of bound pronouns, one according to which quantifiers, names, and pronouns can all bind pronouns, indeed one which makes semantic sense of the seemingly transitive binding in a sentence like (5) and at the same time respects the Binding Theory:

- (5) John says that he loves his wife.  
 (5')  $[John\ x] (x\ \text{says that } ([he\ y: y=x] ([his\ z: z\ \text{wife } y] (y\ \text{loves } z))))$ .

(ii) If one is determined to resist the idea that names are quantifiers, all is not lost. For there are reasons for thinking that the relevant abstraction emerges in another way. On an account of predication that has much to commend it independently, an intransitive V, say 'snores' amounts to a one-place predicate  $\lambda x(x\ \text{snores})$ .<sup>136</sup> An intransitive verb, say, 'loves', amounts to a two-place predicate  $\lambda y(\lambda x(x\ \text{loves } y))$ . On such an account,  $\alpha$ 's binding  $\beta$  amounts to  $\alpha$ 's merging with a  $\lambda$ -predicate whose operator binds  $\beta$ .<sup>137</sup> Schematic logical forms of (1) and (2) are given by (1'') and (2''):

- (1) John loves his wife.  
 (1'')  $John(\lambda x(x\ \text{loves } x\ \text{'s wife}))$   
 (2) every man loves his wife  
 (2'')  $[every\ y: man\ y](y(\lambda x(x\ \text{loves } x\ \text{'s wife})))$ .

Spelling out the finessed 'his' we get (1''') and (2'''):

- (1''')  $John(\lambda x([his\ z: z\ \text{wife } x](x\ \text{loves } z))$   
 (2''')  $[every\ y: man\ y](y(\lambda x([his\ z: z\ \text{wife } x](x\ \text{loves } z))))$ .

<sup>135</sup> If Montague's approach is taken, the quantifier would be regarded as standing for a higher-order property: 'every man' would stand for that property true of just those properties every man has; 'John' would stand for that property true of just those properties John has.

<sup>136</sup> One reason for espousing this approach to predication is that it appears to explain what is going on in sentences containing conjoined VPs without seeing such sentences as sentence conjunctions. For example, (i) is understood as (ii):

- (i) John snores and sleepwalks  
 (ii)  $John(\lambda x(x\ \text{snores} \bullet x\ \text{sleepwalks}))$ .

To better mirror English surface forms, here and throughout I put the lambda predicate after the term it applied to.

<sup>137</sup> Heim and Kratzer (1998), Reuland (2001a,b), Reinhart (2000), Salmon (1986, 1992), Soames (1990, 1994).

We now seem to be in a position to say something about locality. The VPs ‘loves him’ and ‘loves himself’ are both formed by merging the verb ‘loves’ with a DP. On the assumption that economy drives languages to contain devices whose job is to register mandatory binding, there will surely be pressure on such devices to be locally bound. The V ‘loves’ is understood as  $\lambda y(\lambda x(x \text{ loves } y))$ , and it is implicitly asking the DP with which it combines to form a VP how  $y$  stands with respect to  $x$ . The reflexive replies, ‘ $y=x$ ’, thus terminating discussion—the question cannot arise again for *that* reflexive later in the building process—meaning that the reflexive is locally bound. (In an alternative framework, the reflexive might be understood as  $\lambda P\lambda xPx$ .) By contrast, the non-reflexive ‘him’ replies, ‘ $y \neq x$ ’, leaving the matter open.

On this view, ‘himself’ might be seen as the reflexive form of the accusative ‘him’. The absence of a reflexive form of the nominative ‘he’ would now be explained: the subject of S is outside the scope of the  $\lambda$ -operator introducing the VP of S. This leaves us with ‘his’, about which I shall have much to say later—to foreshadow, it is the surface form of distinct reflexive and non-reflexive pronouns.

Relatedly, we must now take on the matter of subject-orientation. In many languages, reflexives can be bound only by subject expressions. This is exactly as one would expect on the current proposal. But of course English is not such a language as we saw earlier with (6), which contrasts in this respect with the Icelandic (6’):

- (6) John<sup>1</sup> told Paul<sup>2</sup> a lot about himself<sub>1/2</sub>  
 (6’) *Jón*<sup>1</sup> *sagði Páli*<sup>2</sup> *frá (sjálfum) sér*<sub>1/\*2</sub>.

Let us hold off on this matter until we have more data to work with.<sup>138</sup>

(iii) Can we now dispense with talk of *de jure* co-reference in an account of pronominal anaphora? Examples like (7) might suggest not:

- (7) John loves his wife. Yesterday I saw him buying her roses again. He spends a lot of money on flowers. His wife is very lucky.

On the proposed unification we have no account of the apparently cross-sentential, unbound anaphora exemplified here. We can say that the first occurrence of /his/ is

<sup>138</sup> Corresponding to the simplex reflexives *sig* (ACC), *sér* (DAT), and *sín* (GEN) are complex counterparts *sjálfan sig*, *sjálfum sér*, and *sjálfs sín*, hence the parenthetical *sjálfum* in (6). Simplex reflexives get up to tricks their well-behaved complex counterparts do not (see below). This is also the case in many other languages—and not just the Scandinavian or even Germanic languages—and it is widely held that this is directly connected to the fact that the simplex reflexives have fewer ‘ $\phi$  features’, which include things like person, number, and gender, which at least in principle could make them less restrictive in application. Contrast English ‘himself’, which can only be used of males, and Icelandic *sig* which is not so restricted.



bound by ‘John’; but then we appear to be stuck. The occurrences of /him/ and /he/, and the second occurrence of /his/ are not within the scope of ‘John’ so they are not bound; yet they can certainly be used to refer to John. But does this mean they are *anaphoric* on ‘John’ in any sense relevant to grammatical theory? Is there anything problematic involved in saying (following Kripke and Lewis in one context, and Lasnik and Bach in another) that John has been raised to sufficient salience by the use of ‘John’ in the first sentence of (6), that he is a reasonable target for indexical uses of /he/, /him/, and /his/ in the subsequent sentences? If not, then it seems we have not yet found a good reason for thinking that co-reference plays a role in a theory of anaphora.

But are we absolutely certain we can dispense with the notion of *de jure* co-reference in cases of *c-command* anaphora? Here matters are tricky, as we can see by looking at an example of linguistic ellipsis. The literature on this form of ellipsis is vast and consensus is not easy to find, but traditionally linguistic ellipsis is subject to a stringent parallelism condition on form and interpretation.<sup>139</sup> ‘A constituent may be deleted at PF only if it is a copy of another constituent at LF’, as Heim and Kratzer (1998: 250) put it, moreover a copy *interpreted in the same way*. We may as well call this *recoverability* as it is basically today’s analogue of what used to be known as the *recoverability of deletion* (see Chomsky (1964)), itself basically a consequence of Katz and Postal’s (1964) hypothesis that transformational rules do not affect meaning.<sup>140</sup> Compare (8) and (9):

- (8) John loves himself, and Paul does too
- (9) John loves his wife, and Paul does too.

There is a single reading of (8). But there are two quite distinct readings of (9):

---

<sup>139</sup> For recent, user-friendly discussions of linguistic ellipsis, see Heim and Kratzer (1998) and May (2002).

<sup>140</sup> As Heim and Kratzer note, ‘A main insight of work on ellipsis in the 1970s was that the relevant level had to be one which was fully disambiguated with respect to interpretation’ (1997: 250). As Bach (1981), Larson and Segal (1995), and Ostertag (1998, 1999) have warned, it is important to distinguish two quite different uses of the words ‘ellipsis’ and ‘elliptical’ in the philosophy of language. The first, a *pragmatic* notion, is found in explicit talk of elliptical *uses* or elliptical *utterances* of sentences, by Quine (1940), Sellars (1954), Kripke (1977), Bach (1981), Neale (1990), and others in their respective appraisals of Russell (1905). The second, a *grammatical* notion, is found in explicit talk of elliptical *sentences* by Strawson (1954) and Stanley (2002), which corresponds to linguistic ellipsis in generative linguistics. There can be only embarrassment in store today for those who run the two together. It is somewhat surprising that Strawson (1954), who in 1950 impressed upon us so forcefully the distinction between *sentences* and *utterances of sentences* in his critique of Russell, misses it completely in his critique of Sellars’s (1954) defence of Russell, despite the fact that Sellars himself (i) respects the sentence-utterance distinction, (ii) consistently talks of elliptical *utterances*, and (iii) makes it very clear this is what he is talking about in connection with utterances of descriptions. As a result, Strawson’s objections, parroted by Stanley (2002), are stillborn. For detailed textual analysis and a discussion of the issues, see Neale (2004).

(9')  $Paul(\lambda x([his\ z: z\ wife-of\ x](x\ loves\ z)))$

(9'')  $Paul(\lambda x([his\ z: z\ wife-of\ John](x\ loves\ z)))$ .

The former is usually called the *sloppy* reading, the latter the *strict* reading. The sloppy reading is readily explained on the assumption that 'his' is bound by 'John' in the first clause of (8). On that assumption, the first clause predicates  $\lambda x([his\ z: z\ wife-of\ x](x\ loves\ z))$  of John, the second predicates it of Paul.<sup>141</sup>

But what about the strict reading? There would appear to be two ways to go here. The first involves abandoning the dream of dispensing with *de jure* co-reference altogether in a final theory of anaphora: 'John loves his wife' is ambiguous between two distinct anaphoric readings, one on which 'John' binds 'his', yielding the sloppy reading of (9) as before, the other on which 'his' is a device of *de jure* co-reference, referentially dependent upon 'John', yielding the strict reading. However, this proposal might be thought to be undermined by the absence of a strict reading of (8), which demonstrates that 'John loves himself' has only a bound reading (if 'himself' could be used as a device of *de jure* co-reference, referentially dependent upon 'John', a strict reading of (8) should be available). On the other hand, the contrast might be construed as merely illustrating an important difference between reflexives and non-reflexives: the former permit only bound readings.

The alternative proposal involves saying there is only one anaphoric use of 'his', the bound use, and then explicating the strict reading of (9) in terms of an *indexical* use of 'his', one upon which it is used to refer to John, who has been made salient by the use of 'John' (thus taking a leaf out of Lasnik's book). This seems to me the correct analysis, but it raises an interesting issue. Principle B prevents 'him' being *bound* by 'John' in (10):

(10) John loves him.

That is, Principle B prevents (10) being read as (10'):

(10')  $John(\lambda x(x\ loves\ x))$ .

But aren't we now forced to posit a principle that prevents a speaker from using 'him' indexically in an utterance of (10) to refer to John, to block the reading whose truth conditions we might represent using (10'')?

(10'')  $John(\lambda x(x\ loves\ John))$ .

---

<sup>141</sup> Keenan (1971), Sag (1976), Williams (1977), Heim and Kratzer (1998), Reinhart (1983), Grodzinsky and Reinhart (1993). The common predication in singular and quantified pairs seems to explain the fact that we understand and immediately see the validity of the sloppy (i):

(i) Every Englishman loves his wife, but John doesn't, so John can't be English.

This is far from obvious. First, I might utter (10) whilst pointing at a man I do *not* take to be John but *do* take to be someone John loves, unaware that it is actually John. At most it would seem we need a principle that prevents a speaker from using ‘him’ indexically in (10) to refer to someone he takes to be the referent of his use of ‘John’. But even this seems too strong. Suppose you and I are sitting at a sidewalk café discussing John. You say to me, ‘He loves no-one’. I am about to reply, ‘Untrue. John loves himself’, when I notice John across the street and instead utter, ‘Untrue. John loves him’, pointing at John, knowing that you will recognize him immediately.<sup>142</sup>

On this account, only indexical readings of the pronouns are possible in (11)-(13), as they do not lie within the scope of ‘John’:

- (11) /his wife loves John/
- (12) /the woman he married loves John/
- (13) /the woman who married him loves John/.

This seems to me a good result. First, only indexical readings of the pronouns are available in the quantified counterparts of (11)-(13):

- (11') /his wife loves every man/
- (12') /the woman he married loves every man/
- (13') /the woman who married him loves every man/.

If we are serious about giving a unified account of what is going on in quantified-singular pairs like (1) and (2), we should be just as serious about giving a unified account of what is going on in the quantified-singular pairs (11')/(11), (12')/(12), and (13)/(13')

For each of (11)-(13) there is certainly a reading upon which the pronoun and ‘John’ are co-referential. But that is no threat to the theory at hand: the occurrences of /his/, /he/, and /him/ in (11)-(13) can be used indexically to refer to whoever ‘John’ is being used to refer to, John being as salient as any other potential target of an indexical occurrence of a masculine pronoun (at least by the time the name ‘John’ is uttered). With utterances of the quantified examples, by contrast, there is

---

<sup>142</sup> Reinhart (1983, 2000), Grodzinsky and Reinhart (1993), Reinhart and Reuland (1993), and Reuland 2001*b*) have explored various formulations of some sort of interpretive condition on (roughly) intrasentential co-reference. The underlying thought is if a given message can be conveyed by two LFs differing only in whether a particular pronoun  $\beta$  is construed as bound by  $\alpha$  or as co-referential with  $\alpha$ , the former is to be strongly preferred. An alternative has been explored by Heim (1993).

no corresponding individual to target (i.e. no individual who is also the intended referent of the relevant DP, which is quantificational).<sup>143</sup>

If this general framework is right, we need to be careful with our words again, particularly ‘pragmatic’. Identifying whether *A* is using /his/ in the bound or in the indexical way *is* a pragmatic matter: although *B*’s knowledge of word meaning and syntax tell *B* that *A* must be using /his/ in either the bound or indexical way, it does not provide *B* with enough information to establish *which one A intends* on any particular occasion. At the same time, *if A* is using /his/ in the bound way, then the pronoun’s interpretation is *fixed relative to that use*: if *B* happens to know that *A* is using /his/ in the bound way—suppose, for the sake of argument, that *A* reveals this to *B*—there is no *further* pragmatic work for *B* to do with respect to /his/: its interpretation is now secured. By contrast, *if A* is avowedly using /his/ in the indexical way, there is still some pragmatic work for *B* to do: *H* needs to establish whether *A* is using /his/ to refer to, say, John, or Paul, or George or Ringo.

All of this serves to indicate once again that we must pause to reflect when we come across claims that the interpretation of pronouns, or of certain pronouns, or

---

<sup>143</sup> Do we have here the germ of an argument for a semantically distinct referential reading of definite descriptions? Larson and Segal (1995) appear to think so. Here is a version of the argument they mount, cast in the terminology of the present chapter. Consider the following:

- (i) /his wife loves the man I talked to/
- (ii) /the woman he married loves the man I talked to/
- (iii) /the woman who married him loves the man I talked to/.

Binding is out of the question—witness (11’)-(13’). But seemingly contrary to expectations, these examples pattern with the *singular* examples (11)-(13), not with the quantified examples (11’)-(13’), in admitting of readings upon which the direct object and the pronoun buried in the subject DP seem to co-refer. How is this to be explained? Answer: the pronoun must be functioning as a referring expression that inherits its referent from the direct object, just as in (11)-(13) where the direct object is a referring expression. But in that case the descriptions in (i)-(iii) must also be functioning as referring expressions.

This argument, is essentially a syntactically elaborate version of the Argument from Anaphora presented by Strawson (1952) for a referential reading of descriptions, and as such appears to fall short of its target for reasons given by Ludlow and Neale (1991) and Neale (1990) using an observation made by Kripke (1977) and Lewis (1979). Using a description referentially—no one doubts referential *usage*, only its semantic significance—can make a particular individual salient (just as using a name can), rendering that individual a natural target for an occurrence of a pronoun used freely. The Russellian does not deny that the pronouns in (i)-(iii) may be used to refer to the same individual the speaker is referring to using the descriptions; at the same time, he agrees with the syntactician that these pronouns are *not bound by* those descriptions. There is no contradiction here, indeed no tension whatsoever. If the antecedent description is used *non-referentially*, as it might be in a particular of, say, (iv)

- (iv) his prime minister walks behind the monarch

plausibly it is descriptive (see below). That we might be on the right track seems to be supported by the following contrast:

- (v) his wife loves a man I was just talking to
- (vi) his wife loves a man.

It is easier to use the richer ‘a man I was just talking to’ referentially here than it is to use ‘a man’ so.

of certain occurrences of pronouns is a ‘pragmatic’ matter. All sorts of claims, some correct, others false or misguided, may lurk beneath the surface of such loose remarks.

Does the fact that /his/ admits of distinct bound and indexical uses mean that it is *ambiguous*, and that (1) and (2) are correspondingly ambiguous?<sup>144</sup> The fear of being assailed for postulating ambiguities has driven philosophers to heroic lengths in preserving unitary semantic analyses.<sup>145</sup> But ambiguity is a tricky notion: some forms (by whatever fancy name) are seemingly less expensive than others; some may result in theoretical simplification elsewhere; and some may make more sense if seen from the perspective of more than one language, as we shall see.

### 13. Bound and Free

Suppose we were to find languages in which there is no unique translation of either (1) or (2) because quite unrelated pronouns—rather than mere morphological variants—were used depending upon whether a bound or indexical use of /his/ is intended:

- (1) /John loves his wife/
- (2) /every man loves his wife/.

And suppose these languages contained only one translation of (3), containing the indexical pronoun where English has ‘his’:

- (3) /his wife loves John/.

A native English speaker might say that these languages make a pointless lexical distinction. But a native speaker of a Scandinavian language (Danish, Faroese, Icelandic, Norwegian, or Swedish), in which there appears to be a *lexical* distinction between genitive and possessive pronouns (the latter being reflexive), might say that the English /his/ is ambiguous, albeit in a systematic way. Translating (1) or (2) into Icelandic, for example, requires fixing whether or not /his/ is bound by the subject expression:

- (1') *Jón<sup>1</sup> elskar konuna sína<sub>1</sub>*  
John loves wife-the self<sup>s</sup>-FEM+ACC+SG

<sup>144</sup> The question is somewhat reminiscent of the question whether /the/ is ambiguous because definite descriptions have distinct attributive and referential uses. My own view is that /the/ is not ambiguous (at least not in a way that the attributive-referential distinction brings out).

<sup>145</sup> Classic cases of this involve the interpretation of /the/, /a/, /and/, /or/, /if/, /because/, /know/, and /see/.

(‘John<sup>1</sup> loves his<sub>1</sub> wife’).

- (2') [sérhver maður]<sup>1</sup> elskar konuna sína<sub>1</sub>  
 every man loves wife-the self's-FEM+ACC+SG  
 (‘[every man]<sup>1</sup> loves his<sub>1</sub> wife’).

The word *sína* is a reflexive possessive (or possessive reflexive), feminine, accusative, and singular to agree with *konuna*, the noun it qualifies (which is why I have rendered it as *self's*, rather than *his*).<sup>146</sup>

But if the English pronoun is understood indexically, indicated here with a subscripted arrow, (1) and (2) must be translated as follows:<sup>147</sup>

- (1'') *Jón elskar konuna hans*  
 John loves wife-the he-MASC+GEN+SG.  
 (‘John loves his<sub>↓</sub> wife’).
- (2'') *sérhver maður elskar konuna hans*  
 every man loves wife-the he-MASC+GEN+SG.  
 (‘every man loves his<sub>↓</sub> wife’).

Here *hans* is the simple genitive, which (unlike *sína*) occurs in the masculine and enters into no agreement relations whatsoever with *konuna*.<sup>148</sup> The important interpretive point, is that *sína* is always bound and hence always within the scope of its antecedent.<sup>149</sup>

<sup>146</sup> The Icelandic definite article typically takes the form of a suffix added to the noun, its origins in a free-standing definite article which is rarely encountered in ordinary talk today). Without the suffix the noun is typically understood as indefinite. So, for example, in the nominative (used primarily for the subject of a verb) the feminine noun *kona* (‘woman’ or ‘wife’) becomes *konan* (‘the woman’ or ‘the wife’) when definite. In the accusative (used primarily for the direct object of a verb but also with some prepositions) *konu* becomes *konuna* when made definite, as in the example discussed above. Apart from certain standardized exceptions, the counterparts of English possessive descriptions are formed using the definite rather than the indefinite form of the noun. There are two main types of exception. The first is where the noun to which the possessive is attached expresses a close family relation, with the notable exception of *kona* (which can translate ‘woman’ as well as ‘wife’). The second will be explained once the difference between possessive and genitive pronouns in Icelandic has been set out.

<sup>147</sup> The arrow notation is borrowed from Evans (1980), but it may be viewed here as a formal non-numerical subscript, the presence of which blocks assignment of a numerical subscript.

<sup>148</sup> The genitive is also used with proper names (and with common nouns) to indicate possession. Thus the accusative form of ‘John’s wife’ will be *konu Jóns*—that is, unlike in cases involving possessive pronouns (*konuna sína*) and cases involving genitive pronouns (*konuna hans*), in cases involving genitive names (*konu Jóns*), the suffix for the definite article is *not* added to the noun; thus the second type of exception mentioned a moment ago.

<sup>149</sup> Linguists who have investigated Icelandic often point to special contexts in which a simple reflexive may appear despite lying strictly outside the scope of its purported antecedent. (See (e.g.) Maling (1984), Þráinsson (1991), Reuland and Sigurjónsdóttir (1997).) Consider (i):

On the assumption that anaphoric readings are marked *syntactically*, what exactly do we say about the English (1) and (2)? Given the data from Icelandic, one might consider associating two distinct LFs with each, one in which /his/ is a bound pronoun (subscripted with a numeral) and one in which it is an indexical:

- (4)  $[_{\text{DP}}\text{every man}]^1 [_{\text{S}} e_1^1 \text{ loves his}_1 \text{ wife}]$   
 (5)  $[_{\text{DP}} \text{ every man}]^1 [_{\text{S}} e_1 \text{ loves his}_\downarrow \text{ wife}]$ .

But on the proposal outlined earlier, once these LFs are specified more precisely it is tempting to say that ‘his’ is unambiguous: it is an unambiguous D, part of DP, part of the NP of which needs a value, and that value can be fixed through binding or, if unbound, indexically:

- (4′)  $[_{\text{DP}}\text{every man}]^1 [_{\text{S}} [_{\text{DP}}\text{his}^2 [_{\text{NP}}\text{wife } e_1]^2] [_{\text{S}} e_1 \text{ loves } e_2]]$   
 (5′)  $[_{\text{DP}}\text{every man}]^1 [_{\text{S}} [_{\text{DP}}\text{his}^2 [_{\text{NP}}\text{wife } e_\downarrow]^2] [_{\text{S}} e_1 \text{ loves } e_2]]$ .

- 
- (i)  $[_{\text{DP}}\text{trú Jóns}^1] \text{ virðist vera að þú elskir sig}_1/\text{konuna sína}_1$   
 $[_{\text{DP}}\text{belief John's}] \text{ seems be that you love-SUBJ self/wife-the self's}$   
 (‘John’s<sup>1</sup> belief seems to be that you love him<sub>1</sub>/his<sub>1</sub> wife’)

The following features appear to be crucial to these examples: (a) the subject expression must express something like a perspective or centre of consciousness of the individual referred to by *Jón*; (b) the verb must be in the subjunctive. In a switch to my perspective all is lost:

- (ii)  $[_{\text{DP}}\text{trú Jóns}^1] \text{ fær mig til að halda að þú elskir sig}_{*1}/\text{konuna sína}_{*1}$   
 $[_{\text{DP}}\text{belief John's}] \text{ leads me to think that you love-SUBJ self/wife-the self's}$   
 (‘John’s<sup>1</sup> belief leads me to think that you love him<sub>\*1</sub>/his<sub>\*1</sub> wife’).

Such facts suggest to some linguists that these reflexives are *logophors* rather than *anaphors*. A pronoun  $\beta$  is *logophoric* if (i) some antecedent expression  $\alpha$  has been used to refer to some individual  $x$ , and (ii)  $\beta$  is now used to refer to  $x$  from  $x$ ’s own perspective. (The perspective manifests itself in the stretch of discourse containing  $\alpha$  and  $\beta$ —for example, the speaker or writer might be reporting or more generally representing  $x$ ’s speech, thoughts, feelings, or perceptions.) (Logophors were brought to the attention of linguists by Hagège (1974), to whom the term appears to be due, and Clements (1975) in their investigations of certain African languages.) In connection with Icelandic the logophor hypothesis is the hypothesis that in certain subjunctive constructions that crucially involve something like the perspective of some individual, a simple reflexive need not occur within the scope of its antecedent. Indeed it is arguable that such a reflexive may have *no overt antecedent whatsoever*, as in the following example discussed by Sigurðsson (1990):

- (iii) *María var alltaf svo andstyggileg.*  
 Mary was always so nasty.  
*Pegar Ólafur kæmi segði hún sér áreiðanlega að fara...*  
 When Olaf would come-SUBJ, she would certainly tell self to leave  
 [self ≠ Olaf, but the person whose perspective or thought is being represented].

On the other hand, the contrast between (i) and (ii) may bear some similarity to the contrast in English between the near c-command cases (iv) and (v):

- (iv) [Bush’s opinion] is that pictures of himself in uniform all over town will capture votes.  
 (v) [Bush’s campaign] requires pictures of himself in uniform to be put up all over town.

(iv) is certainly more acceptable than (v) to my ear (in my dialect ‘him’ must replace ‘himself’ in both, but (v) seems more forgivable). For interesting recent discussion of the issues here, see Reuland (2001a).

## 14. Pragmatic Anaphora?

There is no shortage of English sentences in which a pronoun  $\beta$  appears to be anaphoric on an expression  $\alpha$  (singular or quantified) yet does not lie within  $\alpha$ 's scope and hence cannot, on our assumptions, be bound by  $\alpha$ . Let us use subscripts in parentheses if the (purportedly) anaphoric connection is not one of binding:

- (1) John<sup>1</sup> is cleaning the house. His<sub>(1)</sub> wife returns today
- (2) If John<sup>1</sup> is late, Mary will be upset with him<sub>(1)</sub>

It would seem sufficient to say that the use of the name 'John' in utterances of (1) and (2) renders some individual salient enough to be a natural target for indexical uses of the pronouns 'his' and 'him'. If we insist on using the words 'anaphora' and 'anaphoric' here, let us preface them with a qualifier like 'seeming' or 'unbound' or 'discourse' or 'pragmatic'. (The Icelandic translations of (1) and (2) comport with this idea, requiring unbound *hans* for /his/, and unbound *hann* for /him/, respectively). Continuing with Postal's idea that 'his wife' and 'him' are descriptions, the idea would be that [<sub>DP</sub>him [<sub>NPe</sub>]] in (2), for example, is understood as [*him*  $x_1$ :  $x_1 = x_2$ ] with  $x_2$  free to refer indexically to John.

What about cases in which a pronoun is seemingly anaphoric on a quantified DP outside whose scope it lies, for example (3) and (4)?

- (3) John bought [only one donkey]<sup>1</sup> and Paul vaccinated it<sub>(1)</sub>
- (4) If John buys [just one donkey]<sup>1</sup> then he pays cash for it<sub>(1)</sub>

As Evans (1977) points out, construing the pronouns in these examples as variables bound by the quantified DPs upon which they appear to be anaphoric, by giving the quantifiers large scope, yields the wrong results. Someone who utters (3), for example, would not be claiming that only one donkey satisfies *John bought  $x$  and Paul vaccinated  $x$* . For that claim is consistent with John buying two donkeys, while the claim made by uttering the original conjunction in (3) is not.

If genuine grammatical anaphora involves binding, what are we to say about these discourse pronouns? We have construed a pronoun seemingly anaphoric on but not bound by a *name* as an indexical co-referential with the name. What is the analogue of this where the seeming antecedent is a quantified DP? A number of philosophers and linguists have argued that that the pronouns in question *go proxy for descriptions* and hence are themselves quantified DPs (assuming, Russell's Theory of Descriptions) just like their purported antecedents.<sup>150</sup> On such an

<sup>150</sup> Cooper (1979), Davies (1981), Ludlow and Neale (1991), Neale (1990, 1994, 2004). In a different vein, see also Elbourne (2001, 2002). Such theories have their origins in Evans's (1977) theory of E-type pronouns, which several of the aforementioned people, as well as Soames (1989), see as both confused and empirically flawed. According to Evans, the pronouns in (3) and (4)—and, indeed, all



account, the occurrence of the pronoun ‘it’ is understood as if it were an occurrence of the description ‘the donkey John bought’; in (2) it is understood as if it were an occurrence of ‘the donkey John buys’. Let us call pronouns that go proxy for descriptions in this way *D-type* pronouns. Since descriptions may enter into scope relations with other expressions, the D-type proposal can explain certain ambiguities. For example, (5) appears to have two readings according as /he/, understood as going proxy for ‘the man who assaulted the queen last night’, is read with small or large scope in the second sentence:<sup>151</sup>

- (5) A man assaulted the queen last night. The police think he’s an escaped convict.

The D-type proposal appears to be successful with examples like (6) and (7) involving what is known variously as ‘relativization’, ‘covariation’, or ‘implicit binding’:

- (6) every villager owns [a donkey]<sup>1</sup> and feeds it<sub>(1)</sub> at night  
 (7) every villager who bought [a donkey]<sup>1</sup> vaccinated it<sub>(1)</sub>

The pronoun /it/ is naturally understood as going proxy for the description /the donkey he bought/, read with smaller scope than the subject quantifier, capturing the implicit binding of the pronoun /he/ in the description for which ‘it’ goes proxy.<sup>152</sup> Obviously the D-type proposal harmonizes with the general idea that

others apparently anaphoric on quantified expressions that do not bind them—form a natural group in *having their references fixed rigidly by descriptions*, singular or plural as the case may be. In (3), for example, /it/ has its reference fixed by the ‘the donkey John bought’; and in (4), /it/ has its reference fixed by ‘the donkey John buys.’ (On Evans’s account, the larger group to which E-type pronouns belongs also includes descriptive names such as ‘Julius’, introduced by a description such as ‘the man who invented the zip’. See Evans (1982, 1985) for discussion.)

<sup>151</sup> This ambiguity seems beyond the reach of Evans’s E-type theory. In the same vein, consider (i):

- (i) [Mary<sup>1</sup> wants [<sub>S</sub>PRO<sub>1</sub> to marry [a wealthy man]<sup>2</sup>]]. He<sub>(2)</sub> must be a millionaire.

The first sentence in (i) may be read *de dicto* (as the Russellian would put it, the scope of ‘a wealthy man’ may be small, just the embedded sentence). Moreover, the pronoun ‘he’ in the second clause can be discourse-anaphoric on ‘a wealthy man’. But as Karttunen (1976) notes, on this *de dicto* reading the modal expression ‘must’ must be present for the discourse anaphora to work. Compare (i) with (ii):

- (ii) [Mary<sup>1</sup> wants [<sub>S</sub>PRO<sub>1</sub> to marry [a wealthy man]<sup>2</sup>]]. He<sub>(2)</sub>’s a millionaire.

In (ii) it is not possible to get the *de dicto* reading for the antecedent clause if ‘he’ is discourse-anaphoric on ‘a wealthy man’. The contrast between (i) and (ii) is explicable on the assumption that ‘he’ is interpreted as if it were the description ‘the man Mary marries’, which may take small scope with respect to the modal ‘must’ in (i) (an infelicitous existence implication results if it is interpreted with large scope). In (ii) on the other hand, since there is no modal operator with respect to which the pronoun can be understood with small scope, the sentence has no felicitous reading when the antecedent clause is read *de dicto*.

<sup>152</sup> On Evans’s E-type account, /it/ has its reference fixed in (6) and (7) by the description /the donkey he owns/, the interpretation of ‘he’ covarying with objects that ‘villager’ is true of in (6) and objects that ‘villager who bought a donkey’ is true of in (7). So although the pronouns in (6) and (7)

third person pronouns are descriptions. There is no explicit commitment in the proposal itself, however, to the view that whereas the PFs of (6) and (7) contain ‘it’ their LFs contain the fleshed out definite description ‘the donkey he owns’; but that is certainly one option that might be explored.<sup>153</sup>

With the distinction between bound and ‘pragmatic’ anaphora in mind, let us reflect for a moment on the nature of the distinction between, on the one hand ‘his<sub>k</sub>’ and the Icelandic *sína* (etc.), and on the other ‘his<sub>i</sub>’ and the Icelandic *hans*. A donkey sentence containing a relative clause can be used as a diagnostic:

(8) [every man who has [a son]<sup>2</sup>]<sup>1</sup> admires his<sub>1/(2)</sub> wife.

The important feature of (8) is that /his/ is *within* the scope of the subject DP /every man who has a son/ but *outside* the scope of the DP /a son/. When /his/ is anaphoric on the subject DP it is bound; but when it is seemingly anaphoric on /a son/ it is unbound (plausibly going proxy for the relativized description ‘his son’s’). Since Icelandic reflexives (possessive or otherwise) are always bound—as already noted there are seemingly logophoric examples that add complications—it is not surprising that the translation of (8) depends upon whether /his/ is to be read as bound by /every man who has a son/ or as an unbound pronoun seemingly anaphoric on /a son/:

---

are not themselves bound, interpreting them does involve recognizing implicit binding because the interpretation of /he/ in the description that fixes the referent of /it/ must covary with other items. The word ‘implicit’ is vital here, for Evans makes no claim about the existence of a variable in the underlying syntactic structures of (6) or (7) that is bound by the subject quantifier. Certain contortions are required to make Evans’s theory fly here because, to speak very loosely, the description that fixes the reference of the pronoun must be interpreted as within the scope of the subject quantifier in order to get the required variation, and this seems initially strange given that Evans claims E-type pronouns are rigid referring expressions. Soames (1989) notes that when Evans’s formalism is examined, E-type pronouns are effectively equivalent to descriptions with largest scope. On the basis of examples like (6) and (7) Evans ends up making the confused claim that the use of a *binary* quantification [*the*<sub>1</sub>]( $\phi(x)$ ;  $\psi(x)$ ) to represent ‘the  $\phi$  is  $\psi$ ’ is compatible with Russell’s claim that descriptions are incomplete symbols whereas the use of a unary restricted quantification [*the*<sub>1</sub>:  $\phi(x)$ ] $\psi(x)$  is not. See the appendix to Neale (2001) for discussion.

<sup>153</sup> Stanley and Szabó (2000) and Stanley (2000) assume that implicit binding always requires an actual variable at LF, even if anaphoric. Oddly, they do not mention Evans’s (1977) implicit binding or the implicit binding in description-proxy variants of Evans’s account such as those by Davies (1981), Neale (1990), and Ludlow and Neale (1991), mentioned below, all of which are careful not to say they are postulating a variable at LF that the subject quantifier binds. See Elbourne (2001, 2002) for a discussion of the virtues of having descriptive material in the LF. If one holds the truth-evaluable conception of LFs, then perhaps one would need something like this (although there is a healthy and growing literature on both variable-free logics and variable-free semantics for natural language). There is also no commitment in the D-type proposal to the view that the description for which a particular D-type pronoun goes proxy can be extracted by some automatic procedure from the immediate linguistic context. It does seem clear, however, that there is a default procedure or heuristic that yields extremely good results in very many cases. For discussion, see Neale (1990), Ch. 5.

- (8') [sérhver maður sem á son]<sup>1</sup> elskar konuna sína<sub>1</sub>  
 every man who has son loves wife-the self<sup>o</sup>s-FEM+ACC+SG  
 ('[every man who has [a son]<sup>2</sup>]<sup>1</sup> loves his<sub>1/\*2</sub> wife')
- (8'') [sérhver maður sem á son] elskar konuna hans<sub>(2)</sub>  
 every man who has son loves wife-the his-(MASC+GEN+SG)  
 ('[every man who has [a son]<sup>2</sup>] loves his<sub>(2)</sub> wife').

On the bound, reading, /his wife/ will be rendered as *konuna sína*; on the unbound ('donkey') reading it will be rendered as *konuna hans*.<sup>154</sup>

The situation is parallel when we have a name inside the relative clause:<sup>155</sup>

- (9) [everyone who knows John]<sup>2</sup> pities his<sub>1/(2)</sub> wife
- (9') [allir sem þekkja Jón]<sup>1</sup> vorkenna konunni sinni<sub>1</sub>  
 everyone who knows John pities wife-the self<sup>o</sup>s-FEM+DAT+SG  
 ('[everyone who knows John]<sup>1</sup> pities his<sub>1</sub> wife.')
- (9'') [allir sem þekkja Jón]<sup>2</sup> vorkenna konunni hans<sub>(2)</sub>  
 everyone who knows John<sup>2</sup> pities wife-the his-MASC+GEN+SG  
 ('[everyone who knows John]<sup>1</sup> pities his<sub>(2)</sub> wife.')

<sup>154</sup> It should be noted that an alternative explanation of the data could be provided. Icelandic possessive reflexive pronouns are like non-possessive reflexives in being subject-oriented:

- (i) Jón<sup>1</sup> sagði Páli<sup>2</sup> frá konunni sinni<sub>1/\*2</sub>  
 John told Paul-DAT. about wife-the-DAT self<sup>o</sup>s-DAT  
 ('John told Paul<sup>2</sup> about his<sub>1/\*2</sub> wife')
- (ii) Jón<sup>1</sup> gaf Páli<sup>2</sup> mynd af konunni sinni<sub>1/\*2</sub>  
 John gave Paul-DAT. picture of wife-the self<sup>o</sup>s-DAT  
 ('John<sup>1</sup> gave Paul<sup>2</sup> a picture of his<sub>1/\*2</sub> wife')

And as with non-possessives, the slack is taken up by non-reflexive counterparts (here the genitives) thereby giving rise to apparent Principle B violations:

- (iii) Jón sagði Páli<sup>2</sup> frá konunni hans<sub>2</sub>  
 John told Paul-DAT. about wife-the-DAT his  
 ('John told Paul<sup>2</sup> about his<sub>2</sub> wife')
- (iv) Jón gaf Páli<sup>2</sup> mynd af konunni hans<sub>2</sub>  
 John gave Paul-DAT. picture of wife-the-DAT his  
 ('John gave Paul<sup>2</sup> a picture of his<sub>2</sub> wife')

In (iii) and (iv) *hans* is bound by *Páll*. This demonstrates a point that will be useful in a moment: it is a mistake to think the distinction between reflexive possessives and genitives corresponds exactly to a distinction between bound and unbound. It corresponds, rather, to a distinction between reflexive (hence bound) and non-reflexive (whether bound or free). See below. (It should not be overlooked that issues about the syntax of double object constructions arise in connection with (i)-(iv).)

<sup>155</sup> The verb *vorkenna* ('pity') in (9') and (9'') assigns dative case to *konunni* ('wife') which is why the reflexive possessive pronoun appears in the dative, feminine, singular form *sinni* there. (Unlike the simple reflexive *sig*, which inflects for case only, the possessive reflexive inflects fully to match its noun.) The prepositions *af* and *frá* also assign dative case in examples that will appear later.

<sup>156</sup> It has been suggested to me more than once that the distinction might not involve the structural notion I have been focusing on but a desire to avoid ambiguity. This is easily disproved. First,

Given the distinction between, on the one hand, (8') and (9'), and on the other, (8'') and (9''), and assuming a plausible syntax of sentences in which 'only' modifies a name, we should expect to see the same distinction in translations of (10):<sup>157</sup>

(10) [S[<sub>DP</sub>only [<sub>NP</sub>John]<sup>2</sup>]<sup>1</sup> loves his<sub>1/(2)</sub> wife].

Here, /his/ is within the scope of 'only John' but not within the scope of 'John'. Where /his/ is bound by the DP /only John/, (10) is used to say that John is the only one who satisfies *x loves x's wife*. Where it is used simply to refer indexically to John it is used to say that John is the only one who satisfies *x loves John's wife*. In Icelandic the distinction is indeed reflected in the use of the bound *sína* and the unbound *hans*:

(10') [*aðeins Jón*]<sup>1</sup> *elskar konuna sína*<sub>1</sub>  
only John loves wife-the self's-FEM+ACC+SG  
( '[only John]<sup>1</sup> loves his<sub>1</sub> wife' )

(10'') [*aðeins Jón*]<sup>2</sup> *elskar konuna hans*<sub>(2)</sub>  
only John loves wife-the his-MASC+GEN+SG  
( '[only John]<sup>2</sup> loves his<sub>(2)</sub> wife' ).

As certain double genitives are permitted in Icelandic, however, *konuna hans Jóns* (roughly equivalent to 'his, John's, wife', but not felt to be epenthetic), would almost always be used instead of the simple *konuna hans* in (10'') in order to make one's meaning clearer (this being something of a national obsession).

As one might now expect, the only Icelandic translation of (11) is (11') because the pronoun is not within the scope of the name:

(11) his wife loves John  
(11') *konan hans elskar Jón*

---

consider (i) where there *is* ambiguity, just as in English (I here assume a standard Quine-Evans account of relative pronouns):

(i) *Jón*<sup>1</sup> *vorkennir* [*maninum sem*<sup>2</sup> *elskar konuna sína*<sub>1/2</sub>]  
( 'John<sup>1</sup> pities the man who<sup>2</sup> loves his<sub>1/2</sub> wife' )

Second, consider (ii) and (iii) where there is no relevant ambiguity (assuming no same-sex marriages):

(ii) \**Jón*<sup>1</sup> *er farinn*. *María er að hugga konuna sína*<sub>1</sub>  
( 'John<sup>1</sup> has left. Mary is comforting his<sub>(1)</sub> wife' )

(iii) \**ef Jón*<sup>1</sup> *fer þá mun María hugga konuna sína*<sub>1</sub>  
( 'if John<sup>1</sup> leaves, then Mary will comfort his<sub>(1)</sub> wife' ).

If *sína* is replaced by *hans* in (ii) and (iii), all is fine.

<sup>157</sup> That things should turn out this way was pointed out to me by James Higginbotham.

As in English, intended co-reference is not excluded, but *konan hans* would almost always cede to *konan hans Jóns* in actual speech if co-reference were intended.

Using the distinction between bound and discourse pronouns, we can explain facets of an old chestnut mentioned earlier (for details, see Jacobson (1977, 2000) and Neale (1990)):

(12) [the pilot who shot at it<sub>(2)</sub>]<sup>1</sup> hit [the MiG that was chasing him<sub>1</sub>]<sup>2</sup>

If the subject DP has larger scope in (12), ‘him’ is bound and ‘it’ is D-type pronoun understood as ‘the MiG that was chasing him’. Suppressing some details (in particular, the structure of ‘him’), (12) will be read as follows:

(12') [*the x: pilot x • [the y: MiG y • y chased x](x shot at y)*]  
 ([*the y: MiG y • y chased x](x hit y)*).

(If the object DP has larger scope the situation is reversed: ‘it’ is bound and ‘him’ is D-type, understood as ‘the pilot who shot at it’.) Similarly, in (13) ‘it’ will be a D-type pronoun understood as ‘his paycheck’:

(13) [the man who gave [his paycheck]<sup>2</sup> to his wife]<sup>1</sup> was wiser than  
 [the one who gave it<sub>(2)</sub> to his<sub>1</sub> mistress].  
 (13') [*the x: man x • [his y: y paycheck x]([his z: z wife x](x gave y to z))*]  
 ([*the w: man w • [his y: y paycheck w]([his z: z mistress w]*  
 (*w gave y to z*))] (*x was wiser than w*)).

As noted earlier, the English sentence (14) is ambiguous according as ‘his’ is understood as used to make independent reference to John (the strict reading) or as bound by ‘John’ (the sloppy reading):

(14) John loves his wife, but Paul does not.

The bound/sloppy reading is captured by (15) in Icelandic:

(15) *Jón elskar konuna sína en ekki Páll.*  
 John loves wife-the *self*'s but not Paul  
 (‘John<sup>1</sup> loves his<sub>1</sub> wife but Paul does not’).

To capture the strict reading, however, something like the following must be used:

(16) *Jón elskar konuna sína en Páll elskar hana ekki.*  
 John loves wife-the *self*'s but Paul loves her-ACC. not  
 (‘John<sup>1</sup> loves his<sub>1</sub> wife but Paul does not love her’).

That is, the reading cannot be captured naturally using a VP elliptical sentence.

Two other English examples mentioned earlier raise similar issues. Reasonable Icelandic renderings of (17) and (18) are (17') and (18'), respectively, which seem

to involve not VP ellipsis but something more akin to what is usually called *VP anaphora*, making use of a pronominal expression (*það*, ‘it’):

- (17) if every man loves his wife then clearly John does  
 (17') *efsérhver maður elskar konuna sína, þá gerir Jón það augljóslega líka*  
 (18) every Englishman loves his wife, but John doesn't, so John can't be  
 English  
 (18') *sérhver Englendingur elskar konuna sína, en Jón gerir það ekki, svo Jón getur ekki verið Englendingur.*

## 15. The Binding Theory Revisited

The precise statement of the Binding Theory has exercised linguists for nearly a quarter of a century. I stated a simple version earlier (the subscripts indicate this is our first attempt):

*Principle A<sub>1</sub>*: a reflexive is to be interpreted as bound in its own clause.

*Principle B<sub>1</sub>*: a non-reflexive is not to be interpreted as bound in its own clause.

*Principle C<sub>1</sub>*: a name is not to be interpreted as bound.

This statement served us reasonably well, but it is now time to look at certain problems and potential revisions.

The examples in the previous section do not constitute a knock-down case for scope being the sole structural relation necessary to isolate occurrences of the possessive reflexive in Icelandic. In the English (1) and (2), either ‘John’ or ‘Paul’ may bind the relevant pronoun:

- (1) John<sup>1</sup> told Paul<sup>2</sup> about himself<sub>1/2</sub>  
 (2) John<sup>1</sup> told Paul<sup>2</sup> about his<sub>1/2</sub> wife.

This fact is not mirrored in Icelandic, where reflexives are *subject-oriented*. Only the subject expression *Jón* may bind the simple reflexive *sér* in (1') and the reflexive possessive *sinni* in (2'):

- (1') *Jón<sup>1</sup> sagði Páli<sup>2</sup> frá (sjálfum) sér<sub>1/\*2</sub>*  
 John told Paul-DAT about (EMPH) self-DAT  
 (‘John<sup>1</sup> told Paul<sup>2</sup> about himself<sub>1/\*2</sub>’)  
 (2') *Jón<sup>1</sup> sagði Páli<sup>2</sup> frá konunni sinni<sub>1/\*2</sub>*  
 John told Paul-DAT about wife-the-DAT self'S-FEM+DAT+SG  
 (‘John<sup>1</sup> told Paul<sup>2</sup> about his<sub>1/\*2</sub> wife’).

This contrast between English and Icelandic needs to be explained, and it is *not* explained by the Binding Theory. Similarly the contrast between English (3) and Icelandic (3’):

- (3) [sJohn<sup>1</sup> [VPgave Paul<sup>2</sup> [DPA photograph of himself<sub>1/2</sub>]]].  
 (3’) *Jón<sup>1</sup> gaf Páli<sup>2</sup> ljósmynd af (sjálfum) sér<sub>1/\*2</sub>*  
 John gave Paul-DAT. photograph of (EMPH) self-DAT  
 (‘John<sup>1</sup> gave Paul<sup>2</sup> a photograph of himself<sub>1/\*2</sub>’).<sup>158</sup>

If one wanted to refer back to the referent of *Páli* here one would be forced to use the non-reflexive *honum* (*sjálfum*) in (3’). *Mutatis mutandis*, with ‘his wife’ in place of ‘himself’ in (3), and *konunni sinni* in place of *sig* in (3’). If one wanted to refer back to the referent of *Páli* this time one would be forced to use the non-reflexive *konunni honum*.<sup>159</sup>

As far as English is concerned, the interesting contrast is between (3) and (4):

- (4) John<sup>1</sup> gave [DPPaul<sup>3</sup>’s father]<sup>2</sup> [DPA picture of himself<sub>1/2/\*3</sub>]

‘Paul’ in (3) and ‘Paul’s father’ in (4) both contain ‘himself’ within their scopes, and each occurs in the reflexive’s immediate clause, so Principle A explains why they are potential binders. By contrast, the reflexive is not within the scope of ‘Paul’; in (4), so Principle A explains why ‘Paul’ is *not* a potential binder here.

Following the discussion earlier, we can say that (i) it is a syntactico-semantic fact, known by *A* and *B* if they are speakers of English, that either ‘John’ or ‘Paul’ may bind the pronouns in (1) and (2); and (ii) since *B*’s knowledge of word

<sup>158</sup> (3) and (3’) involve double object constructions *seemingly* of the form [VPV DP DP]. The first DP is called the *indirect* object, and the second the *direct* object. In Icelandic, different verbs may assign different cases to their objects; for example *gefa* (‘give’) assigns dative case to its indirect object, *Páli* in (3’), and accusative case to its direct object, *ljósmynd*. There are two worrying features about the structure [VPV DP DP]: (a) it involves ternary rather than binary branching of the VP node, and (b) the indirect and direct objects lie within one another’s scopes. There are reasons for thinking both of these are problematic and that the syntax of the double object construction is not [VPV DP DP] at all. On the sort of analysis proposed by Larson (1988), for example, (3) and (3’) have structures more like (i) and (i’):

- (i) John [VPgave<sup>i</sup> [VP[DP Paul]<sup>2</sup> [e<sub>i</sub> [DPA photograph of himself<sub>1/2</sub>]]]].  
 (i’) *Jón<sup>1</sup> [VPgaf<sup>i</sup> [VP[DP Páli]<sup>2</sup> [e<sub>i</sub> [DPljósmynd af (sjálfum) sér<sub>1/\*2</sub>]]]]].*

In these structures *e<sub>i</sub>* is an aphonic interpreted via the verb (*gave/gaf*), within whose scope it lies, and which I have adorned with an alphabetically identically superscript. On Larson’s account, we can think of double object verbs as effectively factored into two components (one within the scope of the other). The net result of this analysis is that the direct object is now within the scope of the indirect object, a structural asymmetry that appears to have considerable explanatory value, particularly where binding is concerned. Making the point more firmly requires examining inverted double object constructions and the niceties of Icelandic case; this is not the place for such excursions. I refer the reader to Larson (1988).

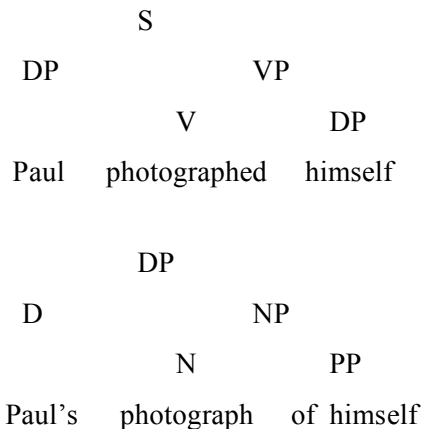
<sup>159</sup> The pronoun *hann* (NOM. ACC.), *honum* (DAT.), *hans* (GEN), is a genuine masculine with constant gender inflection.

meaning and syntax do not provide him with enough information to nail down fully which name *A* intends to be understood as binding /himself/ or /his/ when he utters (1) or (2), *B* must use *non-linguistic* information to discern that intention. In this respect, interpreting /himself/ or /his/ on any particular occasion in which (1) or (2) is uttered is a *pragmatic* matter (constrained by syntax, of course). Furthermore, (i') it is a syntactico-semantic fact, known by *A* and *B* if they are speakers of Icelandic, that only *Jón* may bind *sér* and *sinni* in (1') and (2'); and (ii') since *B*'s knowledge of word meaning and syntax *do* provide him with enough information to establish which name *A* intends to be understood as binding *sér* and *sinni* when he utters (1') or (2'), *B* does *not* need non-linguistic information to discern that intention. And in this respect, interpreting *sér* and *sinni* on any particular occasion in which (1') or (2') is uttered is *not at all* a pragmatic matter.

The notion of *subject* will have to be broadened if the claim that Icelandic reflexives are subject-oriented is to be maintained. Principle A makes correct predictions in cases like the following, where the reflexive and its binder are constituents of a subject DP:

- (5) [<sub>DP</sub>Paul<sup>1</sup>'s photograph of himself<sub>i</sub>] won first prize.  
 (5') [<sub>DP</sub>*ljósmynd Páls<sup>1</sup> af sjálfum sér<sub>i</sub>*] vann fyrstu verðlaun.

Notice *Páls* is not the subject of (5'), yet it still binds *sér*. The explanation for this may well lie in an observation made by Chomsky (1973): from a structural point of view, the possessive determiner 'Paul's' (*Páls*) stands to the entire possessive DP 'Paul's photograph of himself' (*ljósmynd Páls af sjálfum sér*) as a subject DP stands to the whole sentence:



As Chomsky sometimes puts the point, the possessive determiner seems to act as the *subject* of the possessive DP. This idea will assume some importance later.



The fact that reflexives in many language) are subject-oriented is taken by some linguists to be tightly connected to seeming violations of Principle A of the Binding Theory.<sup>160</sup> Let us note first that the following is not a counterexample to Principle A:

(6) John<sup>1</sup> promised Paul to shave himself<sub>1</sub>.

On pretty standard syntactic assumptions, (6) has the following structure,

(6') John<sup>1</sup> promised Paul [<sub>S</sub> *e* to shave himself]

where *e* is an aponic element serving as the subject of the infinitive infinitival clause [<sub>S</sub> *e* to shave himself], which is the reflexive's immediate clause.<sup>161</sup> Moreover, this occurrence of *e* must be bound by 'John', a fact dictated by the syntax and meaning of 'promise', a so-called subject-control or subject-binding verb (*e* is bound by the verb's subject). Hence *e* bears as subscript the same numeral 'John' bears as superscript. Additionally, *e* binds 'himself', which is why it bears as superscript the same numeral 'himself' bears as subscript. In accordance with what has preceded, (6') must shorthand for something more like (6''):<sup>162</sup>

(6'') John<sup>1</sup> promised Paul [<sub>S</sub> [<sub>DP</sub> [<sub>DE</sub>]<sup>2</sup> [<sub>NPe1</sub>]]<sup>2</sup> to shave himself<sub>2</sub>].

[<sub>DE</sub>]<sup>2</sup> is the aponic version of the D 'he', and [<sub>DP</sub> [<sub>DE</sub>]<sup>2</sup> [<sub>NPe1</sub>]]<sup>2</sup> is the aponic version of the whole DP 'he', understood as [*he* *x*<sub>2</sub>: *x*<sub>2</sub>=*x*<sub>1</sub>]. On fairly standard assumptions, the aponic version of the pronoun is required because (i) the verb occurs in the infinitive, (ii) only finite verbs assign case, and (iii) without case a DP may not occur in phonic form.<sup>163</sup>

<sup>160</sup> Chomsky (1986), Pica (1991).

<sup>161</sup> In syntactic theory, this type of phonetically null pronoun is often called PRO (see Chomsky 1981, 1986). I eschew talk of PRO (and also *trace*) here, as the differences between types of aponic DPs are not germane to my main points. It should be stressed, however, that PRO may well have interpretive properties not shared by other pronouns. One difference comes out in comparing the following sentences, due to Grice (1971) (Grice did not use PRO of course):

- (i) I remember PRO biting my nurse
- (ii) I remember my biting my nurse
- (iii) I remember I bit my nurse.

In order for *A* to say something true using (i), *A* must remember performing—not just remember the performance of—a particular nurse-biting. Whilst (ii) and (iii) may certainly be used to say as much, such readings are not forced upon them. The issues here bristle with complexities raised by Castañeda (1966, 1967, 1968), Cherchia (1990), Higginbotham (1990), Lewis (1979), Perry (1979, 2000), and others.

<sup>162</sup> For simplicity I work with an annotated surface form rather than the LF as the main point can be made without setting out the LF.

<sup>163</sup> On 'promise' and other control verbs, see Larson (1991). Also, contrast the following, with clausal adjuncts:

- (i) John<sup>1</sup> learned karate<sup>2</sup> [<sub>S</sub> [<sub>DP</sub> [<sub>DE</sub>]<sup>3</sup> [<sub>NPe1</sub>]]<sup>3</sup> to protect himself<sub>3</sub>]
- (ii) John<sup>1</sup> hired bodyguards<sup>2</sup> [<sub>S</sub> [<sub>DP</sub> [<sub>DE</sub>]<sup>3</sup> [<sub>NPe2</sub>]]<sup>3</sup> to protect him<sub>1</sub>].

Aphonic subjects are not always subject-controlled: ‘persuade’, for example, is an object-controlled verb which dictates that the structure of (7) is something like (7’):

- (7) John persuaded Paul<sup>1</sup> to shave himself<sub>1</sub>.  
 (7’) John persuaded Paul<sup>1</sup> [<sub>S</sub>[<sub>DP</sub>[<sub>De</sub>]<sup>2</sup>[<sub>NP</sub>*e*<sub>1</sub>]]<sup>2</sup> to shave himself<sub>2</sub>].

Again Principle A is observed. But this type of construction leads us straight into an apparent problem posed by Icelandic. In (8), either *Jón* or *Pál* may bind *sig*:

- (8’) [*Jón*<sup>1</sup> *taldi Pál*<sup>2</sup> *á* [<sub>S</sub>[<sub>DP</sub>[<sub>De</sub>]<sup>3</sup>[<sub>NP</sub>*e*<sub>2</sub>]]<sup>3</sup> *að raka sig*<sub>1/3</sub>]]  
 John persuaded Paul-ACC [<sub>S</sub> *x* to shave-INF self-ACC]  
 (‘John<sup>1</sup> persuaded Paul<sup>2</sup> [<sub>S</sub> [<sub>DP</sub>[<sub>De</sub>]<sup>3</sup>[<sub>NP</sub>*e*<sub>2</sub>]]<sup>3</sup> to shave him<sub>1</sub>/himself<sub>3</sub>’).

That is, the grammar of Icelandic appears to allow the binding of *sig* by something outside its immediate clause (if non-indicative), something that Principle A prohibits. We have here what linguists often call *long distance* reflexivization.<sup>164</sup> We find the same situation in (9) (the subjunctive is vital):

- (9) [<sub>S</sub> *Jón*<sup>1</sup> *segir að* [<sub>S</sub> *Páll*<sup>2</sup> *raki sig*<sub>1/2</sub>]]  
 John says that Paul shaves-SUBJ self-ACC  
 (‘John<sup>1</sup> says that Paul<sup>2</sup> shaves him<sub>1</sub>/himself<sub>2</sub>’).

The simplest structural account of the connection between the subject-orientation of Icelandic reflexives and their ‘long-distance’ binding possibilities appears to be wrong: reflexives of both the simplex form (*sig*, *sér*, *sín*) and the morphologically complex form (*sjálfan sig*, *sjálfum sér*, *sjálfs sín*) are subject-oriented, but only simplex reflexives give rise to seeming violations of Principle A.<sup>165</sup>

<sup>164</sup> The label is misleading. Distance *per se* is not the issue, witness the distance between ‘John’ and ‘himself’ in (i):

- (i) John bought a poster of a photo of a sketch of a statue of himself.

The issue is really about the possibility of binding reflexives (and only morphologically simplex reflexives at that) across certain boundaries. Interest in the Binding Theory in connection with Icelandic and ‘long-distance’ phenomena developed largely as a result of a paper by Maling (1984) and then a collection of papers edited by Koster and Reuland (1991), especially the paper by Bráinsson (1991). For earlier discussion, see Bráinsson (1976). For more recent discussion, see Reuland and Sigurjónsdóttir (1997), and the collection edited by Cole *et al.* (2001) especially the paper by Reuland (2001a). I ride roughshod over many points of detail here. Reading Reuland’s papers (after completing the bulk of the present chapter) has made me see various possibilities not reflected here.

<sup>165</sup> There is no hope of sorting out long-distance reflexives here. Complexities involving the proper domain of binding, subject-orientation, intrinsic reflexivity, feature assignment, feature movement, LF movement, VP ellipsis, double-object verbs, factivity, logophoricity, simplicity, and acquisition make the topic daunting. See the interconnected works of Grodzinsky and Reinhart (1993), Reinhart (1983, 2000), Reinhart and Reuland (1991, 1993), Reuland (2001a, 2001b), Reuland and Sigurjónsdóttir (1997), and Sigurjónsdóttir (1992), which are chock-a-block with interesting ideas,

Simplifying somewhat, I want now to set out in a way that philosophers should find immediately accessible, the sorts of problems the Binding Theory as we have stated it must overcome, and run quickly through a series of revisions that have been, or might be, contemplated.

One of the most interesting types of problem arises in connection with *representational* nouns like ‘representation’, ‘picture’, ‘portrait’, ‘sketch’, ‘watercolour’, ‘photograph’, ‘image’, ‘statue’, ‘sculpture’, ‘bust’, ‘bronze’, ‘rendering’, ‘recording’ or ‘description’.<sup>166</sup> Linguists usually call the DPs in which such a noun is the principal noun *picture DPs*. It was noted above that our original Binding Theory correctly predicted the legitimacy of the binding relation in a sentence like (5)/(5’):

- (5) [DP[DPaul<sup>1</sup>’s] [picture of himself<sub>1</sub>]] won first prize.  
 (5’) [DPmynd [DPáls<sup>1</sup>] af sjálfum sér<sub>1</sub>] vann fyrstu verðlaun.

When picture DPs occur in non-subject positions, we encounter problems, however. The unacceptability of one reading of (9) appears to show a revision of Principle A<sub>1</sub> is required:

- (9) [SJohn<sup>1</sup> [VPbought [DP[DPaul<sup>2</sup>’s][NPpicture of himself\*<sub>1/2</sub>]]]].

The problem is that although ‘John’ is a constituent of the smallest sentence containing ‘himself’ (which lies within its scope), something stops it binding it.

Is it the DP boundary? An obvious first shot at revision would have it that S and DP form a unified category for purposes of binding, that a reflexive must be interpreted as bound by an expression in the smallest S or DP containing it—the smallest S or the smallest DP containing it, but the smaller of the two. With future revisions in mind, let [S(β)] be the smallest S properly containing β, let [DP(β)] be the smallest such DP, and let [S/DP(β)] be the smaller of [S(β)] and [DP(β)]. Our second attempt at Principle A is this:

*Principle A<sub>2</sub>*: a reflexive, β, is to be interpreted as bound within [S/DP(β)].<sup>167</sup>

---

but hard-going in places for philosophers. To give a taste (in condensed and simplified form and running a few things together) drawing upon the idea that interpreting a pronoun via binding *as a matter of syntax* is ‘easier’ (in some robust cognitive sense) than interpreting it as merely co-referential with some other expression or as referring to some salient entity, and the idea that the more impoverished the descriptive content of a pronoun (compare *sig* with ‘himself’) the more in need it is of a discourse-central referent, the suggestion is made that a simplex anaphor like *sig* can be interpreted as an independent pronoun with a highly restricted usage.

<sup>166</sup> I suppose one could always cry “Logophors!” here; but let’s see how far we can get without doing so.

<sup>167</sup> Given the hypothesis that pronouns are determiners we need to be careful how we interpret *Principle A<sub>2</sub>* so as not to render it trivially false. If ‘himself’ has the structure [DP<sub>him</sub> [NP<sub>self</sub>]] understood as [*him*  $x_2$ :  $x_2=x_1$ ] in ‘John loves himself’ the bound element (corresponding to  $x_1$ ) is really contained in the NP, in which case ‘John’ does not lie in [S/DP(β)] if β is construed as the

Counterexamples to Principle A<sub>2</sub> are readily produced by replacing the possessive determiner ‘Paul’s’ by a simple determiner:

- (10) [sJohn<sup>1</sup> [v**p**bought [DP[<sub>Da</sub>/this/that][picture of himself<sub>1</sub>]]]].<sup>168</sup>

The contrast between (9) and (10) now has to be explained. It might be thought that in (9) ‘Paul’ must be the binder because it is the *nearer* of two c-commanding DPs in [S(β)] bearing the right number and gender. But (11) and (12) are just as bad as (9) on the relevant indexings:

- (11) [sJohn<sup>1</sup> [v**p**bought [DP[<sub>D</sub>my/your][picture of himself<sub>\*1</sub>]]]]  
 (12) [sJohn<sup>1</sup> [v**p**bought [DP[<sub>D</sub>the museum’s][picture of himself<sub>\*1</sub>]]]].

So an intervening, c-commanding DP of the right number and gender appears not to be *necessary* to block the binding of ‘himself’ by ‘John’ in the structures we are considering; *any* intervening DP will do the trick. And the presence of an intervening, c-commanding DP is not sufficient in *other* structures—at least not in English—witness (13) and (14):

- (13) [sJohn<sup>1</sup> [v**p**told me/Paul<sup>2</sup> about himself<sub>1/2</sub>]]  
 (14) [sJohn<sup>1</sup> [v**p**gave me/Paul<sup>2</sup> [DP[<sub>Da</sub>][NPpicture of himself<sub>1/2</sub>]]]].

This suggests the contrast between, on the one hand, (9), (11), and (12), and on the other, (10) must have *something* to do with the presence in each of the former of a *possessive* DP, one whose determiner is essentially a genitive DP (‘John’s’, ‘that man’s’, ‘every man’s’) occupying the blank position of the schematic DP,

- (15) [DP \_\_\_ [picture of himself]].

A straightforward modification suggests itself. Let [S(DP>β)] be the smallest S containing β and a c-commanding DP. Now [S/DP(DP>β)] will be the smaller of [S(DP>β)] and [DP(DP>β)]. We can try again to formulate a successful Principle A:

*Principle A<sub>3</sub>*: a reflexive, β, is to be interpreted as bound within [S/DP(DP>β)].

The examples we have looked at so far are now accounted for. Principle A<sub>3</sub> also makes the correct predictions in embedded picture constructions, which demonstrate that the absolute distance between ‘John’ and ‘himself’ is immaterial; it is simply the presence of a possessive determiner that blocks binding:

- (16) John<sup>1</sup> bought [DP<sub>a</sub> photograph of [DP<sub>a</sub> sketch of [DP<sub>a</sub> bust of himself<sub>1</sub>]]]]

---

bound element. But on the determiner hypothesis facts about the bound *element* are not what we are interested in, of course, when we contemplate the Binding Theory. We are interested in its facts about the DP ‘himself’, i.e. we are talking about [S/DP(β)] where β is the DP ‘himself’.

<sup>168</sup> Who says you can’t bind into a complex demonstrative?

- (17) John<sup>1</sup> bought [<sub>DP</sub>Mary's/Paul<sup>2</sup>'s/my photograph of [<sub>DPa</sub> sketch of [<sub>DPa</sub> bust of himself<sub>\*1/2</sub>]]]
- (18) John<sup>1</sup> bought [<sub>DPa</sub> photograph of [<sub>DP</sub>Mary's/Paul<sup>2</sup>'s/my sketch of [<sub>DPa</sub> bust of himself<sub>\*1/2</sub>]]]
- (19) John<sup>1</sup> bought [<sub>DPa</sub> photograph of [a sketch of [<sub>DP</sub>Mary'/Paul<sup>2</sup>'s/my bust of himself<sub>\*1/2</sub>]]].

Here is a problem case, however:

- (20) John<sup>1</sup> bought [<sub>DPa</sub> photograph of [<sub>DP</sub>Paul<sup>2</sup> and himself<sub>1/2</sub>]].

In (20) the DP 'Paul' is a constituent of [DP( $\beta$ )], and of [S/DP(DP $\beta$ )] (i.e. of 'Paul and himself'). So (i) why can't 'Paul' bind 'himself' here? and (ii) why is 'John' a legitimate binder of 'himself'?

(i) To answer the first question, it might be suggested that 'Paul' cannot bind 'himself' in (20) because 'himself' c-commands 'John' (just as 'John' c-commands 'himself', 'DP and DP' being structurally symmetric).<sup>169</sup> Alternatively, it might be suggested that 'Paul' *can* bind 'himself' in (20), the reading being merely unlikely except, perhaps, in cases of what we might call *waxwork* anaphora, discussed by Jackendoff (1992). Imagine a photograph of Paul McCartney standing next to a waxwork dummy of himself in Madame Tussaud's; if John Lennon bought the photograph we might use (20) to report the facts. This is not really the right venue for discussing the complexities of such examples, or those that raise related issues, such as the following, due to Nunberg (1977):

- (21) Yeats hated hearing himself read in an English accent.

(ii) If the second response to question (i) is accepted, no answer to the question why 'John' is a legitimate binder of 'himself' seems possible. But the first response might open up an avenue: A binder  $\alpha$  of  $\beta$  must meet a morphological condition and a structural condition. The morphological condition amounts to  $\alpha$ 's agreeing in number and gender with  $\beta$ . We have been assuming that the structural condition amounts to  $\alpha$ 's c-commanding  $\beta$ ; but perhaps this is not quite enough: perhaps we need to add that  $\beta$  must not c-command  $\alpha$ ; and if this is right, we might then invoke the concept not of a DP being a potential binder *simpliciter* (which involves morphology as well as position) but the concept of a DP being a potential binder structurally speaking, i.e. the concept of a DP occupying a position from which it can bind some expression  $\beta$ . In short, we replace "DP $\beta$ " in Principle A<sub>3</sub>

<sup>169</sup> Those who look affectionately upon co-reference versions of Principle C might be attracted to this way of proceeding.

by “DP»β” interpreted as “DP that c-commands but is not c-commanded by β” in order to get closer to the truth of the matter:

*Principle A<sub>4</sub>*: a reflexive, β, is to be interpreted as bound by an expression within [S/DP(DP»β)].

But here’s another problem:

(22) \*John<sup>1</sup> believes that [S[DP<sub>A</sub> picture of himself<sub>1</sub>] is on display.

(22’) \*Jón<sup>1</sup> telur að [S[DP mynd af sjálfum sér<sub>1</sub>] sé á syningu.

John believes that picture of EMPH-self is-SUBJ on display.<sup>170</sup>

Principle *A<sub>4</sub>* does not block this. So what is it about (22)/(22’) that makes it unacceptable on the indicated indexing? Surely it is the S boundary. It looks as though β’s binder must lie inside [S(β)], so we seem to need the following revision:

---

<sup>170</sup> Some speakers claim to find the English (22) acceptable. To my ear it is awful. It must not be confused with the acceptable infinitival form (i) and its acceptable Icelandic translation (i’):

(i) John<sup>1</sup> believes [S[DP<sub>A</sub> picture of himself<sub>1</sub>] to be on display]

(i’) Jón<sup>1</sup> telur [S[DP mynd af sjálfum sér<sub>1</sub>] vera á syningu].

Following Postal (1974), it is not implausible to suppose that (i) and (i’) involve what is called “raising to object.” The subjects of the embedded infinitival clauses are “raised” to become the objects of the main verb. This might be viewed as triggered by the inability of the infinitival to assign case. That the expressions alleged to be raised are really the objects of the main verb is strongly suggested by the assignment of accusative case to pronouns occupying the same positions:

(ii) John<sup>1</sup> believes himself<sub>1</sub> to be intelligent

(iii) John believes him<sub>1</sub> to be intelligent.

When a tensed verb is used, the pronoun appears in the nominative, as a normal subject:

(iv) John<sup>1</sup> believes he<sub>1/4</sub> is intelligent.

Presumably the subject of the embedded clauses in (i) (and in the Icelandic (i’)) is an aponic non-reflexive pronoun bound by the raised DP. And certainly it can itself bind:

(v) John believes himself to have harmed himself.

The binding details are not importantly different from those in (vi)

(vi) John believes he has harmed himself.

Certainly (v) is less appealing than (vi), but that may be because of the repeated ‘himself’ in the former. Chomsky (1973) once suggested explaining examples like (ii) in terms of what he called the Tensed Sentence Condition (TSC), which required a reflexive’s binder to be in the smallest *tensed* clause containing the reflexive. But now the matter of the unacceptability of (vii) would be unexplained:

(vii) John<sup>1</sup> believes [Sme to love himself<sub>1</sub>].

Chomsky (1973) suggested explaining it in terms of what he called the Specified Subject Condition (SSC), which required there to be no intervening subject between the binder and the bound. The occurrence of ‘me’ in (vii) was such a subject and hence blocked the binding of ‘himself’ by ‘John’. If the raising idea just mentioned in the main text is incorrect, then perhaps Principles A and B will have to incorporate the restrictions imposed by the TSC and SSC. With respect to the latter, see the discussion of examples (2) and (2’) in the main text.

*Principle A<sub>5</sub>*: a reflexive,  $\beta$ , is to be interpreted as bound within both  $[S(\beta)]$  and  $[S/DP(DP \gg \beta)]$ .

In effect, this amounts to the following:

*Principle A<sub>6</sub>*: a reflexive,  $\beta$ , is to be interpreted as bound within the smaller of  $[S(\beta)]$  and  $[DP(DP \gg \beta)]$ .

Probably a little work will reveal a problem for Principle A<sub>5</sub>. But the *real* problem with where we're going is this: Principle A (and thereby, presumably principle B) is just getting too messy and complicated to have any real appeal or explanatory value. Our original division between devices satisfying Principle A and those satisfying Principle B dropped out as a trivial consequence of composition, of *merging*. But now we seem to have something considerably more complicated, and it is tempting to think we need to change tack, to see the data as flowing from a simpler Binding Theory together with orthogonal conditions on interpretation.

There are all sorts of ways of proceeding here, and I will mention just one that seems to me not obviously devoid of merit. What feature do S and DP share that might bring them together in a statement of the Binding Theory? Frege thought he saw a connection: Ss and DPs (at least the simplest DPs, names) are both types of *referring* expression. They are Frege's *atoms*, his *building blocks*, the syntactic and semantic categories from which all others are derived (Names and sentences are saturated, they refer to *objects*; predicates are unsaturated, they stand for *concepts*.)

Suppose we follow Frege in viewing DP and S as primitive: a VP is whatever combines with a DP to form an S. We can at least say then say that the Binding Principles we have considered make reference to no *defined* categories. The most serious complications we have seen were induced by genitive determiners formed from expressions of the right syntactic category to engage in binding themselves, expressions that seem to *draw attention to themselves* as potential binders in much the same way as subjects of sentences do. The reason a *subject* does this is obvious: A subject is one *half* of a sentence: it merges with a VP to yield a complete S, and this makes the whole S its scope. This makes subject DPs quite different from object DPs. And of course we have seen that in some languages *only subject* DPs may bind reflexives.<sup>171</sup>

---

<sup>171</sup> English is not such a language, and to that extent it appears more permissive. On the other hand it does not permit the long-distance binding of simplex reflexives we find in, say, German or Icelandic. Perhaps there is *only so much reflexivization* a language can tolerate. Or, more accurately, only so many *rôles* for reflexive forms before we lose any meaningful distinction between local and non-local binders. Having reflexives as victims of *both* non-subject binding and long-distance binding in the same system just put too much pressure on it. Or, if a logophoric account of apparently long-distance binding is plausible, having reflexives as victims of *both* non-subject binding and rampant

Chomsky (1970, 1973) suggested that genitive DPs as well as Ss may be viewed as having subjects. And later, when exploring refinements of the Binding Theory, Chomsky (1986) suggested that statements of Principle A might have to make reference to subjects in this sense. On such an account, ‘Paul’ will be the *subject* of the DP ‘Paul’s photograph of himself’. If the subject of DP hypothesis were wrong, Icelandic, being subject-oriented in respect of reflexive-binding should not permit what English allows in (2):

(2) [sJohn<sup>1</sup> [v**P**bought [DP[DPaul<sup>2</sup>’s] [NPpicture of himself<sub>\*1/2</sub>]]]]

But as we saw earlier, *it does* permit it:

(2’) *Jón keypti* [DPmynd [DPáls<sup>2</sup>] af (sjálfum) sér<sub>1/2</sub>]  
 John bought picture Paul-GEN of (emph-DAT) self-DAT  
 (‘John<sup>1</sup> bought Paul<sup>2</sup>’s picture of himself<sub>1/2</sub>’).

It’s as if *Páls* is acting as the subject of the object DP. So let us try another tack. Where XP is any category, we might consider the following revision:

*Principle A<sub>7</sub>*: a reflexive,  $\beta$ , is to be interpreted as bound by a DP within [XP(SUBJECT,  $\beta$ )].

What’s nice about this conceptually—forget any coverage problems for a minute—is that it seems to get us back where we started: *merging is what creates scope*. The primary role that DP plays in creating a sentence—merging with VP—is a role it replicates in the creation of further DPs: it merges with NP in the same way it merges with VP, except it does so in its possessive form and functions syntactically as a D. The act of merging creates a subject’s scope, and a reflexive within its scope is *bindable*, thus closing off the binding question.

## 16. Ambiguity Revisited

The Binding Theory is meant to explain among other things the seemingly complementary distribution of reflexives and non-reflexives in respect of binding possibilities and requirements. Where Principle A requires a reflexive (more generally an ‘anaphor’) to have a binder inside some local domain (to be specified), Principle B requires a non-reflexive (more generally a pronominal) *not* to have a binder in the same domain.

Despite evident problems, the theory seems to go some way towards explaining a wealth of data about the behaviour of reflexive and non-reflexive pronouns

---

logophoric usage may just put too much pressure on the system. It’s as if reflexive forms are being asked to do not inconsistent jobs but *too many jobs at once*—and something has to give.



across languages. If, as we have been assuming, ‘he’, ‘him’, ‘his wife’, and ‘himself’ have the same general syntactic structure,  $[_{DP}DET[_{NP}N]]$ , examples (1) and (2) illustrate a simple dilemma:

- (1) John<sup>1</sup> loves his<sub>1</sub> wife.
- (2) John<sup>1</sup> says that Paul loves his<sub>1</sub> wife.

Since we are ignoring the DP boundaries on the target DPs themselves, sentence (1) appears to show that ‘his<sub>k</sub>’ (or at least ‘his NP’) behaves like the reflexive ‘himself<sub>k</sub>’, and unlike the non-reflexives ‘he<sub>k</sub>’ and ‘him<sub>k</sub>’: it is bound by something in  $[S/DP(\beta)]$  (assume, for the sake of argument, this is the local domain). So either ‘his<sub>k</sub>’ is a reflexive and falls under Principle A, or else it is not a reflexive and Principle B needs to be modified. But sentence (2) appears to show that ‘his<sub>k</sub>’ behaves like the *non*-reflexives ‘he<sub>k</sub>’ and ‘him<sub>k</sub>’, and unlike the reflexive ‘himself<sub>k</sub>’: it is *not* bound by something in  $[S/DP(\beta)]$ ; it is bound by ‘John’. (Moreover, an occurrence of ‘his<sub>k</sub>’ could be bound by ‘Paul’ in the same environment.) So either ‘his<sub>k</sub>’ is a non-reflexive and falls under Principle B, or else ‘his<sub>k</sub>’ is a reflexive and Principle A needs to be modified.

One way out of this dilemma would be to say that we have misunderstood the nature of possessive DPs, that ‘his’ does not stand to its phonic NP as ‘he’ and ‘him’ do to their aphonic NPs. Aphonicity itself cannot be the key, however, because the absolute possessive DP ‘his’ can be used wherever ‘his wife’ can be used, for example in (1) and (2) above. Sorting this out would require having a well thought out account of the structure of possessive DPs, and that I do not have. So let me pursue another idea.

By way of softening up the ground, notice the dilemma recurs with co-ordinated DPs:

- (3) John<sup>1</sup> took photos of his<sub>1</sub> wife and himself<sub>1</sub> on the yacht,
- (4) John<sup>1</sup> knows that Paul took photos of his<sub>1</sub> wife and him<sub>1</sub> on the yacht.

If (3) is grammatical, it would seem ‘his<sub>k</sub>’ and ‘himself<sub>k</sub>’ must fall under the same principle of the Binding Theory. If (4) is grammatical, so must ‘his<sub>k</sub>’ and ‘him<sub>k</sub>’; but how can this be if ‘himself<sub>k</sub>’ and ‘him<sub>k</sub>’ fall under different principles and are in complementary distribution for binding purposes?

The solution I will sketch sees a distinction between two distinct pronouns /his<sub>k</sub>/, a reflexive that occurs in (1) and (3), and a non-reflexive that occurs in (2) and (4). Icelandic, as we have seen, makes a *lexical* distinction along what seem like these lines:

- (5) a. *Jón<sup>1</sup> elskar konuna sína<sub>1</sub>*
- b. *Jón elskar konuna hans*

In (5a), *sína* is a reflexive possessive that has to be understood as bound by *Jón* to satisfy Principle A. In (5b), by contrast, *hans* is a genitive pronoun that *cannot* be understood as bound by *Jón* (it occurs in the masculine to indicate the gender of its referent). It is a non-reflexive pronoun satisfying Principle B. I want now to look at English through Icelandic eyes, as it were.

If our aim is uncover very general principles of the language faculty that help explain language acquisition and data across distinct languages, we should at least consider the possibility of a cross-linguistic distinction between reflexive and non-reflexive possessives. For it might turn out that the best overall theory of pronouns posits the existence of precisely such a distinction in English, /his<sub>k</sub>/ being the superficial manifestation of *two* distinct bound possessives, one reflexive the other non-reflexive, which we can represent as ‘<sub>+</sub>his<sub>k</sub>’ and ‘<sub>-r</sub>his<sub>k</sub>’ respectively (the former functioning as what could easily have materialized as the non-existent form ‘himself’s’).<sup>172</sup> On such an account, we would find the former in (6) and the latter in (7):

- (6) John knows that [<sub>S</sub> Paul<sup>2</sup> loves <sub>+</sub>his<sub>2</sub> wife]  
 (7) John<sup>1</sup> knows that [<sub>S</sub> Paul loves <sub>-r</sub>his<sub>1</sub> wife].

Initially this might seem extravagant. But notice that we may need a reflexive possessive ‘their’ in (8a) to explain the fact that it appears in the same syntactic position as the reciprocal in (8b):

- (8) a. [the senators]<sup>1</sup> love their<sub>1</sub> secretaries  
       b. [the senators]<sup>1</sup> love each other<sub>1</sub>’s secretaries.

In the first instance, we would need to distinguish two sentences, (9a) and (9b) below:

- (9) a. John<sup>1</sup> loves <sub>+</sub>his<sub>1</sub> wife  
       b. John loves his<sub>1</sub> wife.

(9a) contains an occurrence of the bound reflexive possessive ‘<sub>+</sub>his<sub>k</sub>’ (corresponding to the occurrence of *sína* in the Icelandic (8a)) whilst (9b) contains an occurrence of the independent possessive ‘his<sub>1</sub>’ (corresponding to the occurrence of *hans* in the Icelandic (8b)).

Now comes the more interesting part. (10) also has two translations in Icelandic:

<sup>172</sup> The possessive form ‘his own’ is sometimes said to be the English reflexive possessive, but this does not appear to be correct, witness the following old favourite (given to me by Stanley Peters) in which ‘his own’ does not have a c-commanding antecedent:

- (i) The judge doesn’t believe him. His own lawyer doesn’t believe him.

(10) John knows that Paul loves his wife.

And in (11a), *sína* can only be understood as bound by *Páll*:

(11) a. *Jón*<sup>1</sup> *veit að Páll*<sup>2</sup> *elskar konuna sína*<sup>\*1/2</sup>

John knows that Paul loves wife-the self's  
(‘John<sup>1</sup> knows that Paul<sup>2</sup> loves his<sup>\*1/2</sup> wife’).

b. *Jón*<sup>1</sup> *veit að Páll* *elskar konuna hans*<sub>1</sub>

John knows that Paul loves wife-the his  
(‘John<sup>1</sup> knows that Paul<sup>2</sup> loves his<sub>1/\*2</sub> wife’).<sup>173</sup>

Now (10) and (11) are more interesting than their predecessors. Earlier it was convenient to think of the distinction between *sína* and using *hans* as corresponding to a distinction between ‘his<sub>k</sub>’ and ‘his<sub>↓</sub>’. As far as simple examples like (5) and (9) were concerned, we made no serious error in seeing things that way. But (10) and (11) demonstrate that this is not actually the right way to think about the difference between *sína* and *hans*. It corresponds not to a difference between bound and unbound but to a distinction between reflexive (hence bound) and non-reflexive (whether bound or free). For if *sína* is used as in (11a), it *must* be read as bound by *Páll*. If one wants a pronoun bound by *Jón* in this structure, one must use (11b) just as one must if one wants an indexical pronoun. So the right distinction in Icelandic appears to be between a reflexive possessive (*sína*) and a non-reflexive possessive (*hans*). The former must be bound; the latter may, but need not, be.

All of this suggests the following mapping between English and Icelandic for a possessive pronoun (applying to a male, but attached to a feminine noun in the accusative):

---

<sup>173</sup> The verb *vita* (‘know’) is factive so we find an indicative embedded clause after *veit*. If a non-factive verb such as *segja* (‘say’) is used the embedded clause may be subjunctive and *Jón* may indeed bind *sína* (just as it may bind *sig* in relevantly parallel environments):

(i) *Jón*<sup>1</sup> *segir að Páll*<sup>2</sup> *elski konuna sína*<sub>1/2</sub>  
John says that Paul loves-SUBJ wife-the self's  
(‘John<sup>1</sup> says Paul<sup>2</sup> loves his<sub>1/2</sub> wife’).

Replacing *sína* in (i) by the non-reflexive genitive pronoun *hans* produces a reading in which some third party is the expected referent. However, although *hans* cannot be construed as bound by *Jón* in the variant, it can be understood as used indexically to refer to John, used in (perhaps) an awkward way of resolving an ambiguity by ‘brute force’ as it were. (I owe this wording to Ólafur Páll Jónsson.)

Possessive			
Bound		Unbound	
Reflexive	Non-reflexive		
$+_r\text{his}_k$	$-_r\text{his}_k$	$\text{his}_\downarrow$	English
$\text{sína}_k$	$\text{hans}_k$	$\text{hans}_\downarrow$	Icelandic

Or, if we prefer to focus on reflexivity rather than binding:

Possessive			
Reflexive	Non-reflexive		
	Bound	Unbound	
$+_r\text{his}_k$	$-_r\text{his}_k$	$\text{his}_\downarrow$	English
$\text{sína}_k$	$\text{hans}_k$	$\text{hans}_\downarrow$	Icelandic

So with an eye to generality, perhaps we should say that both languages contain *three* distinct pronouns here. In English all three coincide phonetically, but in Icelandic only two do.<sup>174</sup> It would be interesting to know whether there are languages in which only the counterparts of ‘ $+_r\text{his}_k$ ’ and ‘ $-_r\text{his}_k$ ’ coincide phonetically—I doubt there are languages in which only the counterparts of ‘ $+_r\text{his}_k$ ’ and ‘ $\text{his}_\downarrow$ ’ coincide—and languages in which all three are phonetically distinct.

On this model, we can see the two bound possessive’s ‘ $+_r\text{his}_k$ ’ and ‘ $-_r\text{his}_k$ ’ at work in (13a) and (13b), respectively:

(13) a. John knows that [<sub>S</sub> Paul<sup>2</sup> loves  $+_r\text{his}_2$  wife]

(13) b. John<sup>1</sup> knows that [<sub>S</sub> Paul loves  $-_r\text{his}_1$  wife].

In (13a), the pronoun is bound from within [DP/S(β)]; in (13b), by contrast, the pronoun is *not* bound from within [DP/S(β)]. Given that ‘ $+_r\text{his}_k$ ’ is a reflexive and

<sup>174</sup> Things go in the other direction too. The third person contrast between the reflexive *sig* and the non-reflexive *hann* is absent in the first and second persons: *mig* is used for *myself* and *me*, and *þig* for *yourself* and *you* (in the accusative):

- (i) *þú rakar þig*  
(‘you shave yourself’)
- (ii) *Jón rakar þig*  
(‘John shaves you’)

Grammar books often say that Icelandic *lacks* first and second person reflexives, that the first and second person personal pronouns are used *instead*. This is not a helpful way of describing matters. For one thing, it would mean that Icelandic first and second person pronouns trivially violate Principle B: (i) contains a non-reflexive that is bound by something in its immediate clause (*þú*). A better characterization is this: in the first and second person, reflexive and non-reflexive pronouns share the same form. That is, Icelandic contains two pronouns *þig*, one of which is reflexive, the other of which is not.

‘<sub>-r</sub>his<sub>k</sub>’ a non-reflexive, this is as it should be—at least if the Binding Theory is on the right track. Both of these must be distinguished, of course, from (13c):

(13) c. John knows that [<sub>S</sub> Paul loves his<sub>↓</sub> wife].

Similarly, we ought really to go back to (11b) and break it down in (11b) and (11c):

(11) b. *Jón<sup>1</sup> veit að* [<sub>S</sub> *Páll<sup>2</sup> elskar konuna hans<sub>1/\*2</sub>*].

(11) c. *Jón<sup>1</sup> veit að* [<sub>S</sub> *Páll<sup>2</sup> elskar konuna hans<sub>↓</sub>*].

One downside of the proposal is economic: we have postulated a lexico-syntactic ambiguity, in the form of two homophonic bound English possessive pronouns ‘<sub>+r</sub>his<sub>k</sub>’ and ‘<sub>-r</sub>his<sub>k</sub>’. But upon reflection the price does not seem exorbitant. The fact that ‘himself<sub>1</sub>’ and ‘his<sub>1</sub> wife’ may be conjoined to form a DP as in

(14) [every candidate]<sub>1</sub><sup>1</sup> introduced his<sub>1</sub> wife and himself<sub>1</sub> to the governor

seems to show that there must be a possessive /his/ with the same binding properties as the reflexive ‘himself’; and the fact that /he<sub>1</sub>/ and /his<sub>1</sub> wife/ may be conjoined to form a DP, as in

(15) [every candidate]<sub>1</sub><sup>1</sup> told the governor that he<sub>1</sub> and his<sub>1</sub> wife were happily married

seems to show that there must be a possessive /his/ with the same binding properties as the non-reflexive /he/. But we know that /he<sub>1</sub>/ and /himself<sub>1</sub>/ cannot be conjoined in this way because each has the mirror image of the other’s binding properties, as enshrined in usual formulations of Principles A and B.

All of the suggestions about pronouns and binding I have sketched here are based on a few ideas from generative grammar, the philosophy of language, and philosophical logic that seem to intersect in interesting ways, and on a simple and a preliminary comparison of third person pronouns in English and language that is importantly similar but interestingly different. Very likely, once these suggestions are examined by linguists who can bring into the picture more languages, more complex examples, and more powerful machinery, they will require all sorts of revisions or scrapping altogether. The idea that natural languages contain two very different forms of reflexives may have much going for it, members of one type (call it the *sig*-type) being monomorphemic, subject-oriented, not subject to Principle A, and logophorically licensed, members of the other type (call it *himself*-type) being morphologically complex, not subject-oriented, locally bound, and not logophorically licensed. And if so, the facts about English possessives may find a much better explanation when all is revealed. One of the joys of

grammatical theorizing is that since every known theory or claim of genuine interest is frustrated by countless counterexamples that immediately suggest different claims, and one's pleasure often derives from the process of bringing disparate things together or pursuing seemingly batty ideas. Which just reinforces my main point: there is clearly much to be gained by linguists and philosophers working together in this area.

## REFERENCES

- Abney, S. (1987), 'The English Noun Phrase in its Sentential Aspect', Ph.D. diss. (Massachusetts Institute of Technology).
- Austin, J. L. (1962), *How to do Things with Words* (Oxford: Clarendon Press).
- Bach, E. (1970), 'Problominalization', *Linguistic Inquiry*, 1, 121-22.
- Bach, K. (1981), 'Referential/Attributive', *Synthese*, 49: 219-44.
- (1987), *Thought and Reference* (Oxford: Clarendon Press).
- 1994: *Thought and Reference*, 2nd ed., Oxford: Oxford University Press.
- Baker, M. (2001), *The Atoms of Language* (Oxford: Oxford University Press).
- Barwise, J., and R. Cooper (1981), 'Generalized Quantifiers and Natural Language', *Linguistics and Philosophy*, 4, 159-219.
- Barwise, J., and J. Perry (1983), *Situations and Attitudes* (Cambridge, Mass.: MIT Press).
- Bezuidenhout, A. (1997), 'Pragmatics, Semantic Underdetermination and the Referential/Attributive Distinction', *Mind*, 106: 375-409.
- Blackburn, S. (1984), *Spreading the Word: Groundings in the Philosophy of Language* (Oxford: Clarendon Press).
- Blackburn, W. (1988), 'Wettstein on Definite Descriptions,' *Philosophical Studies*. 53: 263-78.
- Blakemore, D. (1987), *Semantic Constraints on Relevance* (Oxford: Blackwell).
- Bresnan, J. (2001), *Lexical-Functional Syntax* (Oxford: Blackwell).
- Cardinaletti, A. (1994), 'On the Internal Structure of Pronominal DPs', *Linguistic Review*, 11: 195-219.
- and M. Starke (1999), 'The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns', in H. van Riemsdijk (ed.), *Clitics in the Languages of Europe*, (Berlin: Mouton), 145-233.

- Carston, R. (1988), 'Implicature, Explicature and Truth-theoretic Semantics' in R. Kempson (ed.), *Mental Representations: The Interface Between Language and Reality* (Cambridge: Cambridge University Press), 155-82.
- (1993), 'Conjunction, Explanation and Relevance', *Lingua*, 90: 27-48.
- (2002), *Thoughts and Utterances* (Oxford: Blackwell).
- Castañeda, H-N. (1966), 'He\*: A Study in the Logic of Self-Consciousness', *Ratio*, 8: 130-57.
- (1967), 'Indicators and Quasi-Indicators', *American Philosophical Quarterly*, 4: 85-100.
- (1968), 'On the Logic of Attributions of Self-Knowledge to Others', *Journal of Philosophy*, 65: 439-56.
- Chierchia, Gennaro (1990), 'Anaphora and Attitudes *De Se*', in R. Bartsch et al. (eds.), *Language in Action*. (Dordrecht: Foris), 1-31.
- Chomsky, N. (1964), *Current Issues in Linguistic Theory* (The Hague: Mouton).
- (1965), *Aspects of the Theory of Syntax*. (Cambridge, Mass.: MIT Press).
- (1970), 'Remarks on Nominalization', in R. Jacobs and P. Rosenbaum (eds.) *Readings in English Transformational Grammar*, (Waltham, Mass.: Ginn and Co.), 184-221; repr. In D. Davidson and G. Harman (eds.) *The Logic of Grammar* (Encino, Calif: Dickenson, 1975), 262-89.
- (1973), 'Conditions on Transformations', in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle* (New York: Holt, Rinehart and Winston), 232-86.
- (1975), *Reflections on Language*. (New York: Pantheon)
- (1976), 'Conditions on Rules of Grammar', *Linguistic Analysis* 2: 303-51.
- (1977). 'On WH-movement', in P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax* (New York: Academic Press), 71-132.
- (1980), *Rules and Representations*. (New York: Columbia University Press).
- (1981), *Lectures on Government and Binding*. (Dordrecht: Foris).
- (1986), *Knowledge of Language: its Nature, Origin, and Use*. (New York: Praeger).
- (1995), *The Minimalist Program*. (Cambridge, Mass.: MIT Press).
- (2000), *New Horizons in the Study of Language and Mind*. (Cambridge: Cambridge University Press).
- (2002), *On Nature and Language* (Cambridge: Cambridge University Press).

- Clements, G. N. (1975), 'The Logophoric Pronoun in Ewe' *Journal of West African Languages* 10: 141-77
- Cole, P., G. Hermon, G., and C.-T. J. Huang (eds.), *Syntax and Semantics 33: Long-Distance Reflexives* (San Diego, CA.: Academic Press).
- Cooper, R. (1979), 'The Interpretation of Pronouns' in F. Heny and H. Schnelle (eds.), *Syntax and Semantics, vol. 10: Selections from the Third Gröningen Round Table*. New York: Academic Press, 61-92.
- Crimmins, M. (1992), *Talk About Belief* (Cambridge, Mass.: MIT Press).
- Crimmins, M. and J. Perry (1989), 'The Prince and the Phone Booth: Reporting Puzzling Beliefs', *Journal of Philosophy*, 86: 685-711.
- Culicover, P. (1976), 'A Constraint on Co-referentiality', *Foundations of Language* 14, 109-18.
- Davidson D. (1986), 'A Nice Derangement of Epitaphs', in E. Lepore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, (Oxford: Blackwell), 433-46.
- Davies, M. (1981), *Meaning, Quantification, Necessity*. London: Routledge and Kegan Paul.
- Elbourne, P. (2001), 'E-Type Anaphora as NP-Deletion', *Natural Language Semantics*, 9, 241-88.
- (2002), 'Situations and Individuals', Ph.D. diss. (Massachusetts Institute of Technology).
- Evans, G. (1977), 'Pronouns, Quantifiers and Relative Clauses (I)', *Canadian Journal of Philosophy* 7, 467-536. (Reprinted in Evans (1985), 76-152.)
- (1980), 'Pronouns', *Linguistic Inquiry* 11, 337-62. (Reprinted in Evans (1985), 214-48.)
- (1982), *The Varieties of Reference*. (Oxford: Clarendon Press)
- (1985), *Collected Papers* (Oxford: Clarendon Press),
- Fiengo, R. and May, R. (1994), *Indices and Identity*. (Cambridge, MA: MIT Press).
- Fodor, J. A. (1975), *The Language of Thought*. (New York: Crowell).
- (1983), *The Modularity of Mind*. (Cambridge, MA: MIT Press).
- (1987), *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. (Cambridge, Mass.: MIT Press).
- (2001), 'Language, Thought, and Compositionality', *Mind & Language* 16: 1-15
- Geach, P. T. (1962), *Reference and Generality*. (Ithaca, NY: Cornell).
- (1972), *Logic Matters*. (Oxford: Blackwell)



- Grayling, A. (1995), 'Perfect Speaker Theory', in J. Hill *et al.* *Reference and Meaning*. Filosofia.
- Grice, H. P. (1957), 'Meaning', *Philosophical Review*, 66, 377-88.
- (1961), 'The Causal Theory of Perception', *Proceedings of the Aristotelian Society*, suppl. vol. 35, 121-52.
- (1971), 'Intention and Uncertainty', *Proceedings of the British Academy*, 263-279.
- (1975a), 'Logic and Conversation', in P. Cole and J. Morgan (eds.), *Syntax and Semantics, vol. 3: Speech Acts*. New York: Academic Press, 41-58. Reprinted in Grice (1989).
- (1975b), 'Method in Philosophical Psychology: From the Banal to the Bizarre', *Proceedings and Addresses of the American Philosophical Association*, 23-53.
- (1989) *Studies in the Ways of Words*. (Cambridge, Mass: Harvard).
- Grodzinsky, Y, and Reinhart T. (1993) 'The Innateness of Binding and Coreference', *Linguistic Inquiry* 24: 69-101.
- Hagege, C (1974) 'Les Pronoms Logophoriques', *Bulletin de la Société de Linguistique de Paris* 69: 287-310.
- Harman, G. (1972), 'Deep Structure as Logical Form', in D. Davidson and G. Harman (eds.), *semantics of Natural Language* (Dordrecht: Reidel), 25-47.
- Heal, J. (1997), 'Indexical Predicates and Their Uses', *Mind*, 106: 619-40
- Heim, I. (1982/1988), 'The Semantics of Definite and Indefinite Noun Phrases' Ph.D. diss. (University of Massachusetts, Amherst, 1982). (Published as *The Semantics of Definite and Indefinite Noun Phrases* (New York: Garland, 1988).
- (1990), 'E-Type Pronouns and Donkey Anaphora', *Linguistics and Philosophy* 13: 137-78.
- (1993), 'Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach', in *MIT Working Papers in Linguistics*, 25: 205-46
- and Kratzer, A. (1998) *Semantics in Generative Grammar*. (Oxford: Blackwell).
- Helke, M. (1971), 'The Grammar of English Reflexives', Ph.D. diss. (Massachusetts Institute of Technology).
- Higginbotham, J. (1980), 'Pronouns and Bound Variables', *Linguistic Inquiry*, 11, 679-708
- (1983a) 'Logical Form, Binding, and Nominals', *Linguistic Inquiry* 14: 395-420.

- (1983b) ‘The Logic of Perceptual Reports: An Extensional Alternative to Situation Semantics,’ *Journal of Philosophy* 80: 100-27.
- (1990), ‘Reference and Control’, in R. Larson, S. Iatridou, U. Lahiri and J. Higginbotham (eds.), *Control and Grammar*. (Dordrecht: Kluwer), 79-108.
- Higginbotham, J., and May, R. (1981) ‘Questions, Quantifiers, and Crossing’, *Linguistic Review* 1: 41-80.
- Huang, C.-T (1995), ‘Logical Form’, in G. Webelhuth (ed.), *Government and Binding Theory and the Minimalist Program*. (Oxford: Blackwell), 125-75.
- Jacobson, P. (1977), ‘The Syntax of Crossing Coreference Sentences.’ Ph.D. diss. (University of California, Berkeley).
- (1999), ‘Towards a Variable-Free Semantics’, *Linguistics and Philosophy*, 22:117-84.
- (2000) ‘Paycheck Pronouns, Bach-Peters Sentences, and Variable-Free Semantics’, *Natural Language Semantics* 8: 77-155.
- Jackendoff, R. (1972), *Semantic Interpretation in generative Grammar*, (Cambridge, Mass.: MIT Press).
- (1992), ‘Madame Tussaud Meets the Binding Theory’. *Natural Language and Linguistic Theory* 10: 1-31.
- Kamp, H. (1981/1984), ‘A Theory of Truth and Semantic Interpretation’, in J. Groenendijk et al. (eds.), *Formal Methods in the Study of Natural Language* (Amsterdam: Amsterdam Centre, 1981), 277–322; repr. in J. Groenendijk et al. (eds.), *Truth, Interpretation, and Information* (Dordrecht: Foris, 1984), 1-43.
- Kaplan, D. (1989), ‘Demonstratives’ and ‘Afterthoughts’, in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. New York: (Oxford University Press) 481-614.
- Karttunen, L. (1969), ‘Pronouns and Variables’, in R. Binnick et al. (eds.), *Papers from the Fifth Regional Meeting of the of the Chicago Linguistic Society* (Chicago: University of Chicago Press), 108-15.
- (1976), ‘Discourse Referents’, in J. McCawley (ed.), *Syntax and Semantics, Vol. 7: Notes from the Linguistic Underground* (New York: Academic Press), 363-85.
- Katz, J. and P. Postal (1964), *An Integrated Theory of Linguistic Descriptions* (Cambridge, Mass.: MIT Press).

- Kayne, R. (2002), 'Pronouns and their Antecedents', in S. Epstein and T. Seely (eds.), *Derivation and Explanation in the Minimalist Program* (Oxford: Blackwell), 133-66.
- Keenan, E. L. (1971), 'Names, Quantifiers, and the Sloppy Identity Problem', *Papers in Linguistics* 4.2, 1-22.
- (2002), 'Explaining the Creation of Reflexive Pronouns in English', in D. Minkova and R. Stockwell (eds.), *Studies in the History of English: A Millennial Perspective* (The Hague: Mouton de Gruyter), 325-55.
- Klima, E. (1964), 'Negation in English', in J. Fodor and J. Katz (eds.), *The Structure of Language*. Englewood Cliffs, NJ: Prentice Hall, 246-323.
- Koster, J. and E. Reuland (eds.) (1991), *Long-Distance Anaphora* (Cambridge: Cambridge University Press).
- Kripke, S. A. (1977), 'Speaker Reference and Semantic Reference', in P. A. French, T. E. Uehling, Jr., and H. K. Wettstein, *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minnesota Press, 6-27.
- (1979), 'A Puzzle about Belief', in A. Margalit (ed.), *Meaning and Use* (Dordrecht: Reidel), 239-83.
- (1980), *Naming and Necessity* (Cambridge, Mass.: Harvard University Press).
- Lakoff, G. (1972), 'Linguistics and Natural Logic', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Dordrecht: Reidel), 545-665.
- Langacker, R. (1966), 'On Pronominalization and the Chain of Command', in D. Reibel and S. Schane (eds.), *Modern Studies in English*. (Englewood Cliffs, N.J.: Prentice-Hall), 160-86.
- Larson, R. (1988), 'On the Double Object Construction', *Linguistic Inquiry*, 19: 335-91.
- (1991), 'Promise and the Theory of Control', *Linguistic Inquiry*, 22: 103-139.
- and S. Cho (1999), 'Temporal Adjectives and the Structure of Possessive DPs', in S. Bird et al. (eds.) *Proceedings of WCCFL 18* (Somerville, Mass.: Cascadilla Press), 299-311.
- and Segal, G. (1995), *Knowledge of Meaning: An Introduction to Semantic Theory*. (Cambridge Mass.: MIT Press).
- Lasnik, H. (1976) 'Remarks on Coreference', *Linguistic Analysis*, 2: 1-22.
- (1989) *Essays on Anaphora* (Dordrecht: Kluwer).

- Lees, R. and E. Klima (1963), 'Rule for English Pronominalization', *Language*, 39: 17-28; repr. in D. A. Reibel and S. A. Schane (eds.), *Modern Studies in English: Readings in Transformational Grammar*. (Englewood Cliffs, NJ: Prentice-Hall, 1969), 145-59.
- Lewis, D. (1972), 'General Semantics', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*. Dordrecht: Reidel, 169-218.
- (1975), 'Adverbs of Quantification', in E. L. Keenan (ed.), *Formal Semantics of Natural Language* (Cambridge: Cambridge University Press), 3-15.
- (1979), 'Scorekeeping in a Language Game', *Journal of Philosophical Logic*, 8: 339-59
- Ludlow, P. (1989), 'Implicit Comparison Classes', *Linguistics and Philosophy* 12, 519-33.
- Ludlow, P. and Neale, S. (1991), 'Indefinite Descriptions: In Defence of Russell', (with P. Ludlow) *Linguistics and Philosophy* 14, 2 (1991), 171-202. (Reprinted in P. Ludlow (ed.), *Readings in the Philosophy of Language*, Cambridge: MIT Press, 1997, 523-55.)
- McCawley, J. (1976), *Grammar and Meaning* (New York: Academic Press).
- Maling, J. (1984), 'Non-Clause Bounded Reflexives in Modern Icelandic,' *Linguistics and Philosophy* 7: 211-41.
- May, R. (1977), 'The Grammar of Quantification', Ph.D. diss. (Massachusetts Institute of Technology).
- (1985), *Logical Form: its Structure and Derivation* (Cambridge, Mass.: MIT Press).
- (2002), 'Ellipsis', *Macmillan Encyclopaedia of Cognitive Science* (London: Macmillan) 1094-1102.
- Montague, R. (1974), *Formal Philosophy: Selected Papers of Richard Montague*, R. Thomason (ed.), (New Haven: Yale University Press).
- (1973), 'The Proper Treatment of Quantification in Ordinary English', in R. Thomason (ed.), *Formal Philosophy: Selected Papers of Richard Montague* (New Haven: Yale University Press, 1974), 247-70.
- Neale, S. (1990), *Descriptions* (Cambridge, Mass.: MIT Press).
- (1993), 'Grammatical Form, Logical Form, and Incomplete Symbols', in A. D. Irvine and G. A. Wedeking (eds.), *Russell and Analytic Philosophy* (Toronto: University of Toronto Press), 97-139; reprinted in G. Ostertag (ed.), *Definite Descriptions: A Reader* (Cambridge, Mass.: MIT Press, 1998), 79-121.

- (1994), 'Logical Form and LF', in Otero, C. (ed.), *Noam Chomsky: Critical Assessments*, (London: Routledge), 788-838.
- (2001), *Facing Facts*. (Oxford: Clarendon Press)
- (2004), 'This, That and the Other', in A. Bezuidenhout and M. Reimer (eds.) *Descriptions and Beyond*. (Oxford: Oxford University Press), 68-181
- (forthcoming a), *Descriptions*, expanded edn. (Oxford: Clarendon Press).
- (forthcoming b), *Linguistic Pragmatism* (Oxford: Clarendon Press).
- (forthcoming c), *Pronouns* (Oxford: Clarendon Press).
- Nunberg, G. (1977). 'The Pragmatics of Reference' Ph.D. diss. (City University of New York).
- Ostertag, G. (1999), 'A Scorekeeping Error', *Philosophical Studies*, 96: 123-46
- Papagragou, A. (1998), 'Modality and Semantic Underdeterminacy', in V. Rouchota, and A. Jucker (eds.), *Current Issues in Relevance Theory* (Amsterdam: John Benjamins), 237-70.
- (2000), *Modality: Issues in the Semantic-Pragmatics Interface* (Oxford: Elsevier Science).
- Partee, B. (1972), 'Opacity, Coreference and Pronouns', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language* (Dordrecht: Reidel), 415-41.
- (1975), 'Deletion and variable binding', in E. Keenan (ed.), *Formal Semantics of Natural Languages* (Cambridge: Cambridge University Press), 16-34.
- (1989), 'Binding Implicit Variables in Quantified Contexts', in *Proceedings of the Chicago Linguistics Society*, 25 (Chicago: University of Chicago Press), 342-65.
- Perry, J. (1979), 'The Essential Indexical', *Noûs* 13: 13-21.
- (1986), 'Thought Without Representation.' *Proceedings of the Aristotelian Society, Supp. Vol.*, 60: 137-51.
- (2000), *The Problem of the Essential Indexical and Other Essays*, expanded edn. (Stanford: CSLI Publications).
- (2001) *Reference and Reflexivity*. (Stanford: CSLI Publications)
- Pica, P. (1987), 'On the Nature of the Reflexivization Cycle', *NELS*, 17 (2): 483-99.
- (1991), 'On the Interaction between Antecedent-Government and Binding: The Case of Long-Distance Reflexivization', in J. Koster, J. and

- E. Reuland (eds.), *Long-Distance Anaphora* (Cambridge: Cambridge University Press), 119-35
- Postal, P. (1966), 'On So-called Pronouns in English', in F. Dineen (ed.), *Report on the Seventeenth Annual Round Table meeting in Washington on Linguistics and Language Studies* (Washington, D.C.: Georgetown University Press), 177-206; repr. in D. A. Reibel and S. A. Schane (eds.), *Modern Studies in English: Readings in Transformational Grammar*. (Englewood Cliffs, NJ: Prentice-Hall, 1969), 201-24.
- (1974), *On Raising* (Cambridge, Mass.: MIT Press).
- Quine, W. V. (1940), *Mathematical Logic*. (Cambridge, Mass.: Harvard),
- (1960), *Word and Object* (Cambridge, Mass.: MIT Press).
- Récanati, F. (1987), "Contextual Dependence and Definite Descriptions." *Proceedings of the Aristotelian Society* 87: 57-73.
- (1989), 'Referential/Attributive: A Contextualist Proposal', *Philosophical Studies*, 56: 217-49. (1993), *Direct Reference: From Language to Thought*, (Oxford: Blackwell).
- (2001), 'What is said', *Synthese*, 128: 75-91.
- Reinhart, T. (1976), *The Syntactic Domain of Anaphora*. Ph.D. diss. (Massachusetts Institute of Technology).
- (1978), 'Syntactic Domains for Semantic Rules', in F. Guenther and S. J. Schmidt (eds.), *Formal Semantics and Pragmatics for Natural Languages*. (Dordrecht: Reidel), 107-30.
- (1983), *Anaphora and Semantic Interpretation*. (London: Croom Helm).
- (2000), 'Strategies of Anaphora Resolution', in H. Bennis et al (eds.), *Interface Strategies*. (Amsterdam: Royal Academy of Arts and Sciences), 295-325.
- and E. Reuland (1993), 'Reflexivity', *Linguistic Inquiry* 24: 657-720.
- Reuland, E. (2001a), 'Anaphors, Logophors and Binding', in P. Cole, G. Hermon, and C.-T. J. Huang (eds.) *Syntax and Semantics 33: Long-Distance Reflexives* (San Diego, CA.: Academic Press), 343-70.
- (2001b), 'Primitives of Binding', *Linguistic Inquiry* 32: 439-92.
- Reuland, E. and Sigurjónsdóttir, S. (1997), 'Long Distance 'Binding' in Icelandic: Syntax or Discourse. in H. J. Bennis, P. Pica, and J. Rooryck (eds.), *Atomism and Binding*. (Dordrecht: Foris), 323-41.
- Ross, J. R. (1967), *Constraints on Variables in Syntax*. Ph.D. diss. (Massachusetts Institute of Technology).
- Rouchota, V. (1992), 'On the Referential/Attributive Distinction', *Lingua*, 87: 137-167.

- (1994), 'On Indefinite Descriptions', *Journal of Linguistics* 30: 441-75.
- Russell, B. (1905), On Denoting. *Mind* 14, 479-493. (Reprinted in R. C. Marsh (ed.), *Logic and Knowledge*. London: George Allen and Unwin, 1956, 41-56.)
- Safir, K. (forthcoming), *The Syntax of Anaphora* (Oxford: Oxford University Press).
- Sag, I. (1976), *Deletion and Logical Form*. Ph.D. diss. (Massachusetts Institute of Technology).
- Salmon, N. (1981), 'Assertion and Incomplete Definite Descriptions', *Philosophical Studies* 42: 37-45.
- (1986), 'Reflexivity', *Notre Dame Journal of Formal Logic* 27: 401-29.
- (1992), 'Reflections of Reflexivity', *Linguistics and Philosophy*, 15: 53-63.
- Saul, J. (2002a), 'Speaker Meaning, What is Said, and What is Implicated'. *Noûs* 36, 228-48
- (2002b), 'What is Said and Psychological Reality: Grice's Project and Relevance Theorists' Criticisms.' *Linguistics and Philosophy* 25, 347-72
- Searle, J. (1969), *Speech Acts* (Cambridge: Cambridge University Press).
- (1979), *Expression and Meaning* (Cambridge: Cambridge University Press).
- (1980). 'The Background of Meaning', in J. R. Searle, F. Kiefer, and M. Bierwisch (eds.), *Speech Act Theory and Pragmatics* (Dordrecht: Reidel).
- Sellars, W. (1954), 'Presupposing', *Philosophical Review*, 63, pp. 197-215.
- Sells, P. (1987), 'Aspects of Logophoricity', *Linguistic Inquiry* 18: 445-79.
- Sigurðsson, H. A. (1990), 'Long Distance Reflexives and Moods in Icelandic', in J. Maling and A. Zaenen (eds.), *Modern Icelandic Syntax* (New York: Academic Press), 309-46.
- Sigurjónsdóttir, S. (1992), 'Binding in Icelandic. Evidence from Language Acquisition', Ph.D. diss. (University of California, Los Angeles).
- Smith, N. V. (1999), *Chomsky: Ideas and Ideals* (Cambridge; Cambridge University Press)
- Soames, S. (1989), Review of Gareth Evans, *Collected Papers*, *Journal of Philosophy* 89: 141-56.
- (1990), 'Pronouns and Propositional Attitudes', *Proceedings of the Aristotelian Society*, 191-212.
- (1994), 'Attitudes and Anaphora', *Philosophical Perspectives*, 9, 251-72.

- Sperber, D., and D. Wilson (1986) *Relevance: Communication and Cognition*. (Oxford: Blackwell).
- (1995), *Relevance: Communication and Cognition*. 2nd edition (Oxford: Blackwell).
- Stanley, J. (2000), 'Context and Logical Form', *Linguistics and Philosophy* 23, 391-434.
- (2002), 'Making it Articulated', *Mind and Language* 17, 149-68,
- Stanley, J. and Z. Szabó, (2000), 'On Quantifier Domain Restriction', *Mind and Language* 15, 219-61.
- Stockwell, R., P. Schachter, and B. Partee (1973), *The Major Syntactic Structures of English* (New York: Holt, Rinehart and Winston).
- Strawson, P. F. (1950), 'On Referring', *Mind*, 59: 320-44.
- (1952), *Introduction to Logical Theory*. (London: Methuen)
- (1954), 'Reply to Mr Sellars', *Philosophical Review* 63, 216-31.
- (1969), 'Meaning and Truth', Inaugural Lecture, Oxford University; repr in *Logico-Linguistic Papers*, (London: Methuen, 1971), 170-89.
- Szabolcsi, A. (1994), 'The Noun Phrase,' in F. Kiefer and K. Kiss (eds.), *Syntax and Semantics 27: The Syntactic Structure of Hungarian*. (New York: Academic Press), 179-275.
- (2003) 'Binding on the Fly: Cross-sentential Anaphora in Variable-free Semantics', in Kruijff and T. Oehrle (eds.), *Resource-Sensitivity in Binding and Anaphora* (Dordrecht: Kluwer), 215-29.
- Wasow, T. (1972), 'Anaphoric Relations in English,' Ph.D. diss. (Massachusetts Institute of Technology); repr. as *Anaphora in Generative Grammar*. (Gent: E. Story-Scientia, 1979).
- Wexler, K., Culicover, P. and Hamburger, H. (1974), *Learning Theoretic Foundations of Linguistic Universals*. Social Science Working Papers, 60. (Irvine, CA: University of California).
- Wiggins, D. (1976), 'Identity, Necessity, and Physicalism', in S. Körner (ed.) *Philosophy of Logic* (Berkeley: University of California Press), 159-82.
- Williams, E. (1977), 'Discourse and Logical Form', *Linguistic Inquiry* 8, 101-39.
- Þráinsson, H. (1976), 'Reflexives and Subjunctives in Icelandic', *Proceedings of NELS 6* (Amherst: University of Massachusetts) 225-39.
- (1990), 'A Semantic Reflexive in Icelandic', in J. Maling and A. Zaenen (eds.), *Modern Icelandic Syntax*. (New York: Academic Press) 289-307.



- (1991), 'Long-Distance Reflexives and the Typology of NPs', in J. Koster, J. and E. Reuland (eds.), *Long-Distance Anaphora* (Cambridge: Cambridge University Press), 49-75.