

1

Towards a Syntactic Model of Interpretation

1.1 Natural Language as a Formal Language?

Knowing a language means being able to segment sounds in that language into units, recover information from those units, and use this information to work out what someone communicating in that language has intended to convey. This statement about what it means to know a language might seem such common sense as to be banal. In fact, however, it is almost universally rejected by linguists. Standardly for linguists, knowing a language is not to be explained directly in terms of having an ability to use it. Rather, the order of explanation is reversed. First, a characterization of a language is given, then what knowing a language is, and, finally, how language can be used, given such knowledge. As will become apparent in due course, this book is to be an implicit plea to return to the common-sense view. But we start from what is familiar.

According to a view which has dominated linguistic methodology right across different theoretical paradigms since the 1950s, natural languages are taken to be definable as a type of formal language. A formal language is devised to describe some object of study, providing the means of giving it a complete characterization. Such a language is usually presented by itemizing some set of basic expressions, from which all and only the well-formed expressions of the language, the sentences, can be induced through recursively applying a set of syntactic rules. Relative to these rules, semantic rules (if defined) characterize the interpretation of the sentence-sized objects in terms of the interpretation of the component parts and their mode of combination.¹

¹By no means all formal systems have accompanying formal definitions of semantics. But a system that lacks formal specification of its semantics may nevertheless be formal in the sense that the statements that can be expressed within the system are articulated to a level of detail such that further statements (indeed the complete set of such further statements) can be PROVABLY derived. Lacking such rules of proof, a formal semantics is essential.

By analogy with formal languages, a natural language may be characterized as a set of sentence strings generated by a grammar – a finite body of rules which assigns all and only such strings a structure (and a phonological characterization), relative to which a semantic interpretation is defined. There are a number of variations on this theme, depending on the form of the syntactic rules, the detailed form of correspondence between syntactic and semantic rules for a given system, and so on; but it is none-the-less the standard conception of natural languages that they are a specialized type of formal language. According to this completely orthodox view (shared at least by categorial grammar and HPSG paradigms; see Morrill 1994, Pollard and Sag 1994), a language is characterized independently of any agent knowing or understanding that language, as some kind of abstract object: knowing a language must then involve standing in a certain propositional attitude – that of ‘knowing’ – to some stored set of rules which comprise the language.

According to a second view which is taken to be in sharp contrast with this view of language,² a language does not have any status as an entity independent of the individual that uses that language – it is exclusively a psychological construct, a body of principles, *I-language*, which a speaker/hearer has internalized (and in this sense ‘knows’). Though this second characterization of language is self-consciously cognitive, the concept of a linguistic system is essentially similar to that of the formal language characterization. This *I-Language* that the speaker knows is, as in the formal language concept of grammar, a body of principles which, in some sense, a speaker has knowledge of – albeit subconscious and unrecoverable. On both concepts of language, the syntax determines what strings of the language are well-formed, assigning them structures relative to which semantic interpretation is fixed.³ On both of these views, the link between knowledge of a language and its application to understanding or speaking is obscure: the only direct evidence for the knowledge said to constitute a grammar of a language is the ability of individual speakers to make judgements of grammaticality, such judgements being supposedly a direct reflection of this subconscious knowledge, accessible independently of the additional constraints and vagaries of performance (see Higginbotham 1987 amongst many others). It is this body of knowledge – said to be describable independent of any application to particular language tasks – which is put to use in performance and understanding, these being said to be not pertinent to the characterization of the knowledge of language itself. The articulation of properties associated with language in use are thus essentially secondary, defined only relative to some prior articulation of a grammar.

²As the Chomskian philosophy of language (Chomsky 1986), this cognitive rewrite of the view that natural languages are formal languages has dominated the field for nearly half a century.

³Note that this is true even in models which have movement, say, to a level of *LF* as the interface level between syntax and semantics, for even with such movement processes, the structures generated are inhabited by expressions of the language, and not by some distinct representation of interpretation assigned to those expressions. As in other fields of theory construction, not all grammars incorporate a formal assignment of interpretation to such syntactic structures: see n.1.

In this book we are going to be introducing a model of language which reflects the common-sense view of language. We shall be setting out a model of natural language understanding in which the development of an interpretation for a string is defined as an incremental left-to-right process of constructing a logical form representing one possible content attributable to the string – in short, a form of parser. The process will involve the top-down development of tree structure representing the logical form. Each node as it is introduced is inhabited by formulae constructed from the words as these become available on a left-right basis, where each such formula is a representation of content which may have been determined in part through a choice from some set of contextually provided values. The only concept of structure is the sequence of partial logical forms, in terms of which this process is defined. As a representational model of interpretation, this framework is one among several (e.g. Jackendoff 1992). What is novel about this formal articulation of the parsing process is that it purports also to provide an explanation of syntactic properties of natural language. Having set out the framework, what we shall argue is that the dynamics of how representations of content are incrementally built up provides a basis for explaining core syntactic phenomena – specifically long-distance dependency and a cluster of related phenomena. The result will be a new twist on the view that the level of representation explanatory of syntactic distribution is a level of semantic representation.⁴ The only level of representation advocated will indeed be that reflecting some assigned interpretation; but that is only the lesser part of the story, for syntactic explanations will make reference not merely to semantically interpretable representations but also to the process of building up such representations on a left-right basis from a sequence of words.

Our demonstration of the fruitfulness of this approach will involve taking a number of current syntactic puzzles and showing how a solution to them emerges naturally from the dynamic perspective we have adopted. Finally, we shall follow up on the consequences of these results, and return to the point we have started from, closing with a reconsideration of the question of what it means to know a language.

1.2 Underspecification in Language Processing

Central to our account will be a concern not merely with **WHAT** interpretation is projected by a natural language expression, where that interpretation is taken as some composite entity; but with **HOW** the components are successively set into a configuration following a left-right sequence and then combined. With this concern uppermost, one particular property of natural languages is immediately apparent. Language processing as a task of establishing interpretation involves manipulating incomplete objects at every stage except at the very end, as shown by the set of partial trees schematically displaying a parsing sequence for (1.1):

⁴This was the view of the generative semanticists arguing in the late 1960s and 1970s that *deep structure* was the level of semantic representation (see McCawley 1968).

(1.1) John upset Mary.

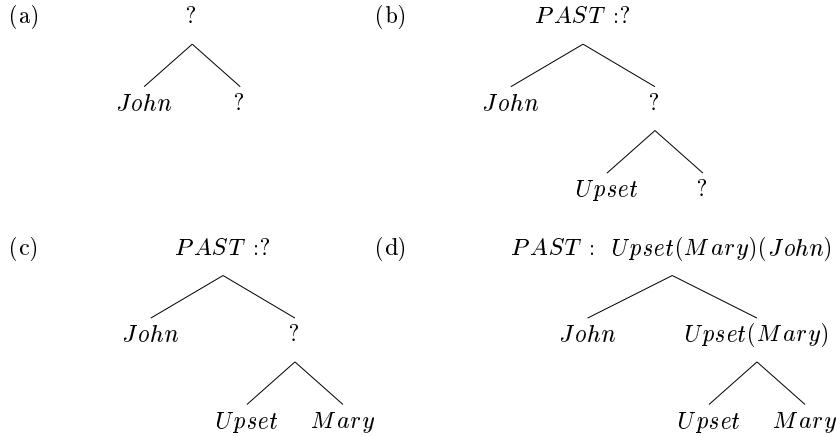


Figure 1.1: The processing of *John upset Mary*

In every tree representation up to the last, we have a partially specified structure for a logical form, in which something is represented as being missing. At step (a), one node is annotated by some logical constant denoting an individual named *John*, and two nodes are not annotated at all, one of which will be subject to further expansion. At (b), a further pair of nodes is introduced, a representation of the relation *Upset* is introduced, but there is as yet no indication of how its second argument is to be saturated, leaving three nodes as yet to be annotated. At (c), the process of unfolding the tree on a top-down basis is complete, the terminal nodes have sub-terms of some logical form, but no interpretation has as yet been assigned to the intermediate nodes. And, finally, at (d), all nodes have been assigned a representation, the interpretation of the root node having duly been compiled by combining the various component parts according to ways directed by the structural configuration. Thus, while the completed tree represents the finally constructed logical form of the string and the mode of combination of its parts, it is the sequence of partial trees which reflects what information is available at each of the intermediate steps. Furthermore, the tree thus constructed represents the logical form for a string, representing one possible interpretation: it is not a syntactic tree (or a parse tree) in the conventional sense of having terminal nodes which are annotated by the words of a string.

This incompleteness of interpretation at all non-final stages of an interpretation process is endemic to the task of parsing. The partial projection of interpretation for a string – simply assuming for the moment that interpretation can legitimately be represented as some sort of annotated tree structure – may involve, as here, partially annotated nodes: some nodes are in some appropriate sense complete, others as yet incomplete, or even yet to be constructed.

The trees displayed in this sequence, however, by no means exhaust the types of incomplete specification which natural language expressions contribute to the build-up of interpretation. On the one hand, this process may also involve taking as input intrinsically incomplete representations of content which get replaced by some antecedent representation already established in the context in which the utterance is being interpreted. This is the phenomenon of anaphora. If we model this phenomenon by defining anaphoric expressions as projecting the trigger that leads to some such process of substitution, then the content lexically projected by such expressions will have to be systematically weaker than any selected interpretation, being rather an indication of the type of substitution which will lead to some appropriate interpretation being established. On the other hand, there may be trees in which individual parts are introduced without specification of how they are related to each other. This phenomenon, we shall argue, is a major part of what has come to be known as the left-periphery problem (see Rizzi 1990), whereby structure associated with left-peripheral constituents in a sentential string may stand in a range of relatively underspecified relations to the structure projected from expressions immediately subsequent to them in the string, at least as initially describable in the parsing process. And this is by no means the end of the list of possible forms of underspecification. A major consideration, later in the book, will be the articulation of such forms of underspecification and how the processes that update these various forms interact.

The need for partial descriptions of tree structure is well recognized in the parsing literature since Marcus (1980) and Marcus et al. (1983), but the trees so described differ from those of the present enterprise in a number of ways. First, standardly, a tree is annotated by words in a string. Secondly, the tree descriptions manipulated in parsing formalisms are defined relative to an independent grammar formalism which defines well-formedness solely in terms of assignment of hierarchical structure to the string. Trees, that is, may be developed into new trees by rules of the grammar, but any such process is invariably defined over (sub)trees which are complete in themselves: there is no concept of a tree in which not all tree relations are specified – this is indeed a distinguishing property of parser formalisms. What will be new in what we introduce is the extension of incomplete specifications from semantics and pragmatics, where they have become familiar through work on anaphora in particular,⁵ to the domain of syntax. Unlike other models, we shall be taking the dynamics of tree growth on a left-to-right basis as the core of an account of intrinsic properties of natural language structure.

On the view we shall put forward, the build-up of interpretation for a string will be projected as a sequence of successively richer descriptions of a logical form represented as a tree structure. The starting point will be a description of a tree as consisting of only a root node, this node decorated merely with a statement

⁵See the postulation of discourse referents in Discourse Representation Theory (DRT) and its variant UDRS (Kamp and Reyle 1993, Reyle 1993, 1996) set up to articulate the process of establishing a DRS, and the postulation of infons, situations, constraints and parameters in Situation Semantics (Barwise and Perry 1983, Cooper 1996).

of the goal – to derive a propositional formula as decoration for that node on the basis of the tree yet to be introduced. The endpoint will be a completely annotated tree structure representing a particular interpretation assigned to that string. The type of example round which much of the subsequent discussion will revolve is that of left-dislocated expressions as in (1.2):

- (1.2) John, Mary is sure that Sue admires.

From a parsing perspective, examples such as this present a particular form of underspecification, in that a language processor, in representing the contribution to the interpretation of the whole made by the left-dislocated expression, is not initially able to identify the role which that expression plays in some propositional structure to be established. Reflecting this informal observation, we will take left-dislocated expressions to annotate a node in the emerging tree which is not initially assigned a fixed tree position, an underspecification which has subsequently to be resolved. This incorporation into the description of natural language structure of underspecified tree relations and their subsequent resolution will turn out to provide a natural basis for explaining the cluster of phenomena associated with long-distance dependency effects across the full array of language types. Such long-distance dependency phenomena are agreed by all to be central to any explanation of the syntactic properties of natural language.⁶ Yet, despite extensive study of long-distance dependency over many years, its characterization has remained surprisingly heterogeneous in all the major frameworks, requiring, as we shall see in due course, a number of different statements with discrete forms of explanation. These include strong versus weak crossover, resumptive pronoun versus gaps, partial movement, *wh-in-situ*, and expletive *wh*, etc. In this new, explicitly dynamic perspective, these phenomena will be seen as falling together in a principled way, while nevertheless allowing systematic cross-language variation.

The sequence of chapters is as follows. In chapters 2 and 3 we introduce the basic framework, with rules for building tree representations providing a logical form for simple clausal sequences, and for long-distance dependency structures. In chapter 4 we give a characterization of the process of building up interpretation for relative clause sequences, with a preliminary indication of how a cross-language typology for these structures might be developed in terms of how information may be copied from one tree structure to another. Setting out this typology will involve surveying restrictive and non-restrictive relatives, head-initial, head-final and even head-internal relatives. In chapter 5 we look at the range of *wh* questions across languages in the light of the dynamic perspective we have been setting out. In chapter 6 we look at the interaction between the processes involved in establishing the interpretation of left-dislocated *wh* expressions and anaphoric expressions – that is, the phenomenon of crossover. What we shall see in these analyses of relative clauses and *wh* questions is that

⁶Recall, back in the 1960s, that it was long-distance dependency phenomena, with their associated island constraints, which it was granted could not be explained in purely denotational terms, with the consequent admission that there were at least some syntactic phenomena whose explanation was irreducibly syntactic (see Partee 1976).

natural typologies emerge as individual language systems are analysed as making available different underspecified descriptions, with consequent variation in the dynamics of their resolution. In chapter 7 we outline ways in which quantification might be reflected in this system. In chapter 8 we step back to consider the general issues raised by this dynamic view. We explore the consequence of this model – a variant of the common-sense view of language – that a grammar of a language provides the architecture within which parse sequences leading to an interpretation can be defined. Indeed, on this view, a string is well-formed if, using all its words in a left-to-right order, it creates at least one logical form. The syntax of natural languages will thus be characterized in terms of incremental building up of structures representing interpretation. The final chapter then provides a more explicit statement of the rules and assumptions set out in chapters 2 and 3.

1.3 The Representational Theory of Mind

The point of departure for this account is the assumption that the process of understanding a natural language string should be defined as a process of incrementally constructing representations. Relatively uncontentious though this assumption may seem, it nevertheless is founded upon some variant of a Representational Theory of Mind, which is taken by many to be controversial (see J.A. Fodor 1981, 1983, 1998, Silvers 1989, Dennett 1987). According to this type of view, assigning interpretation to a signal by a cognitive system necessarily involves pairing the signal with some object (or more abstractly some denotatum of the signal) via the construction of a representation of that object/denotatum. Very approximately, we see a pigeon ahead of us flying across the car, if the given stimulus to the retina causes some cognition-internal representation to be set up which we take to denote that pigeon. The information that can then be derived from such stimuli lies not in the objects themselves, nor in the stimuli, but in the drawing of inferences over such internally constructed representations. As Fodor (1998) expressed it, concepts are essentially mind-dependent. Despite the many controversies surrounding Fodor's particular variant of cognitive psychology, in particular over the semantic properties such a system must have to substantiate what it means to be a representation (see Dennett 1993, Fodor and Lepore 1992), the Representational Theory of Mind is quite generally assumed as a working methodology for cognitive science. As both Fodor and Dennett have put it from opposing stances within the family of representational theories of mind, there does not seem to be any serious alternative, if we are to make any progress in cognitive science. Natural language systems, according to Fodor's own representationalist theory, define a mapping from some relatively low-level representation of the incoming stimulus onto some internal representation within a formal system, over which inference can be syntactically defined, this system of representations being discrete from the natural language itself. It is in this sense that a natural language is what Fodor (1983) calls an *input system*.

A property of all natural language systems that Fodor fails to stress is that

there is a systematic gap between information integral to the natural language expression and the interpretation that is attributed to that expression in a particular context. This is the well-known phenomenon of context dependence, displayed by anaphora, tense, nominal construal and ellipsis, in which the particular interpretation of the expression or structure depends on its processing within the given context:

- (1.3) He ignored her.
- (1.4) Would you like a red apple? Or would you prefer a green one?
- (1.5) I'm going to go home soon. Are you going to, too?

Understanding each of these sentences involves an asymmetry between the information projected by the expressions themselves, *sui generis*, and their interpretation in the context in which they occur. The enrichment of the initial specification to yield some context-specific content has been argued to be constrained by quite general pragmatic considerations such as presumption of relevance (Sperber and Wilson 1986, 1995), and so not a task achieved solely through computations internal to the language input system.⁷ As we shall see in due course, there are various aspects of interpretation which may be incomplete at non-final stages of the parsing process; and the challenge of modelling the task of building up interpretation for natural language strings is to provide a general characterization of the process involved in updating incomplete specifications of interpretation. Preserving the general spirit of the Fodor/Sperber/Wilson Representational Theory of Mind, we shall model all updating of representations as replacement of one description of a partial structure by some description of a richer structure, so that the final representation of content is the result of progressive updates of structure from some starting point. Seen in this light, anaphora resolution and ellipsis construal will be seen as structural processes, which induce a copy procedure from some antecedent representation of content.⁸

In adopting this stance, and moreover anticipating that it might provide a basis for structural properties of natural language systems, the account to be provided will be a departure from current orthodoxies about the nature of natural language interpretation. In particular it involves a departure from the assumption that natural language systems are formal languages whose syntactic properties are defined over strings of words (= sentences) with semantic interpretation defined relative to their syntactic structure. Furthermore, we shall be arguing that the representations resulting from anaphora construal will be the result of interaction between the system-internal operations and quite general pragmatic operations, so that the grammar will systematically underdetermine the full logical forms.

⁷Sperber and Wilson argue, in the face of Fodor's scepticism about the possibility of there ever being a theory of central processing, that interpretation of signals is constrained by a balancing of cognitive effort and inferential effect – a constraint they advocate as the content of the concept of relevance.

⁸See Kamp 1996, which argues for the same conclusion on rather different grounds.

1.4 Pronominal Anaphora: Semantic Problems

It might be argued that any shift to a representational account of the phenomenon of context dependence is a retreat to a purely formal solution that does not provide a semantic explanation. In the past couple of decades, work has been done in particular in Dynamic Semantics, Situation Semantics and Discourse Representation Theory (DRT) on characterizing the relation between an anaphoric expression and its denotational content as part of a general characterization of the context relativity of natural language interpretation (see Kamp 1984, Kamp and Reyle 1993, Groenendijk and Stokhof 1991, Barwise and Perry 1983, and many others). These paradigms vary as to the importance of the partiality integral to the intrinsic content of an anaphoric expression.⁹ However, all these various approaches to the problem of anaphora retain a commitment to characterizing the construal of anaphoric expressions as a property of semantic interpretation, an account which is articulated relative to an independent syntax, within which structural properties of anaphoric expressions are characterized. It should not go unnoticed that this commitment preserves the natural language-as-formal-language methodology, with its assumption that syntax – defined independently of semantics – determines the well-formedness of strings and all their structural properties, leaving the interpretation of the strings to be defined in the light of some such syntactic characterization. Even in DRT, a discourse representation structure itself is a representation of a partial model, and not a level over which syntactic generalizations are defined.

In what follows, we shall be articulating in some detail a representationalist stance on interpretation and, as part of that, a substitutional account of anaphora construal. In the remainder of this chapter, therefore, we provide preliminary evidence for this position, by setting out the problems which face accounts of anaphora in terms of some direct attribution of denotational content to the natural language expression.¹⁰

1.4.1 The Problem of Multiple Ambiguity

The first problem that arises is that pronouns, being interpreted relative to some contextually determined value, will have to be assigned discrete denotational contents according as the context varies, and they will accordingly be

⁹The primary focus in Dynamic Semantics, for example, is solely a characterization of the semantic value of some anaphoric expression relative to what is taken to be a syntactic indexing of the pronoun – without any semantic characterization of the content of a pronoun qua pronoun in the absence of such indexing. Situation Semantics, to the contrary, articulates a whole novel ontology within which a range of partial objects are introduced – situations, infons, parametrized objects, etc. Discourse Representation Theory, preferring a more conservative semantic ontology, articulates an interim construct, the concept of a discourse representation structure, which has a hybrid status both as a form of representation and as a partial model interpretable by embeddability conditions within the total model. In its articulation of an intermediate structural configuration, and only relative to this there being any model-theoretic evaluation, Discourse Representation Theory is much closer to the enterprise we are embarked on here than either Dynamic Semantics or Situation Semantics (see Dekker 1996, Kamp 1996).

¹⁰This section is largely taken from Kempson et al. 1998.

characterized as multiply ambiguous. The challenge, then, for a semantic account of anaphora is whether such an account can provide a unitary base from which to explain the range of interpretations available to a single lexical form. We claim that it cannot. Though pronominal anaphora has been the focus of attention over a considerable period, initiated by Evans (1980), Heim (1982) and Kamp (1984), it is universally assumed that it cannot be reduced to a unitary phenomenon. It is seen as dividing into at least two, if not three, discrete semantic phenomena according as the pronoun is construed as a bound variable, as an indexical referring to some fixed individual, or as a so-called E-type pronoun denoting a witness set of some quantified assertion provided by the antecedent sentence (solutions vary as to how many natural classes are required – see Heim 1990, Groenendijk and Stokhof 1991, Chierchia 1991, Lappin and Francez 1994, Kadmon 1990, van der Does 1996, Neale 1990).¹¹ With these three types (possibly reducible to two), the various forms of intra- and cross-sentential forms of co-reference as in (1.6)–(1.10) can be characterized. However, in addition, anaphoric expressions may have an interpretation established through inference without any direct matching of content, a phenomenon which does not correspond straightforwardly to any of the three categories. The range of types is illustrated in (1.6)–(1.12). (1.6) is an example of a bound variable construal, (1.7) an indexical construal, (1.8) an E-type construal; (1.9) and (1.10) involve intra-sentential co-reference and cross-sentential co-reference; (1.11) is an example of cross-sentential anaphora involving intermediate steps of inference, and the construal of (1.12) involves what has been called *bridging cross-reference* (Clark and Haviland 1977, Clark 1977, Garrod and Sanford 1981):

- (1.6) Every girl worries that she might get pregnant.
- (1.7) She is pregnant.
- (1.8) Most girls passed with distinction. They had worked very hard.
- (1.9) Sue is worried that she might get pregnant.
- (1.10) Sue came in. She was pregnant.
- (1.11) She has either rented a car or she's taken a van. She has to return it on Monday.
- (1.12) The Smiths are nice. He is a doctor.

What is less often pointed out is that these phenomena are replicated right across different forms of anaphoric construal. (1.13)–(1.18) are examples displaying the

¹¹The force of Dynamic Predicate Logic (DPL) has been to define a concept of variable binding which allows some cases of E-type anaphora to fall together with more regular cases (see Groenendijk and Stokhof 1991 and other references). However, this extended concept of variable binding (involving pairs of variables) is not applicable to all cases, leaving a remainder unaccounted for (see van der Does 1996, Kibble 1997). Even DRT, the pioneer in the systematic representation of context dependence and the ability thereby to unify apparently different types of anaphora, presumes a quite different characterization for indexical pronouns (see Kamp and Reyle 1993).

same array of anaphoric effects for definite NPs (in the same relative sequence), and (1.19)–(1.23) and (1.24)–(1.28) exemplify the same phenomena as displayed by the two forms of demonstrative NP (the same putative ambiguity problem also extends to tense construal – see van Eijck and Kamp 1997, Kempson et al. 1998):

- (1.13) Every house I have put on the market I have checked, in order to make sure that the house won't be hard to sell.
- (1.14) The idiot was sick.
- (1.15) Most students were there. The entire group got drunk.
- (1.16) John came in. The idiot had drunk too much.
- (1.17) You can have a car or you can have a van, but the vehicle must be returned on Monday.
- (1.18) The Smiths are nice. The husband is a doctor.
- (1.19) Every day I drink a Coke, I know that later that day I shall have a migraine.
- (1.20) That idiot is sick.
- (1.21) Most people who came early left well before a few people got drunk.
That group were no problem.
- (1.22) John came in with his partner. That idiot had drunk too much.
- (1.23) The Smiths are not as nice as the Jones. That marriage has been a disaster.
- (1.24) Every time I don't take my pills, I think that this time I am better and will not need them.
- (1.25) This idiot is sick.
- (1.26) Most people who came early left well before a few people got drunk.
These drunkards unfortunately spoiled the occasion.
- (1.27) John came in with his partner. This new guy is a disaster.
- (1.28) The Smiths are not as nice as the Jones were. This new couple are never around.

Unless these multiple forms can be explained in terms of some unitary semantic basis, discrete specifications will have to be advocated for every pronoun, every demonstrative determiner, and indeed every tense form. Yet, with an anaphoric expression characterized as a set of discrete lexical items, one would anticipate parametric variation across languages according as one pronoun projects a bound variable interpretation, a discrete pronoun an indexical interpretation, a

third an E-type interpretation. And similarly demonstrative pronouns and tense forms. Yet no such systematic variation exists.¹² The generality of this phenomenon suggests that an optimal solution will involve a unitary base for such lexical anaphora, explaining the semantically divergent forms by some theory of how weakly specified lexical items can give rise to this array of interpretations. For a framework in which lexical content is specified as some fixed denotation (whether as a variable, a witness set or an individual uniquely referred to), this array of interpretations remains problematic.

1.4.2 The Problem of Uniqueness

Turning now to the phenomenon of discourse co-reference, and setting on one side the problem of multiple ambiguity, context dependence is seen quite generally as requiring a semantic characterization in the form of an update mechanism. Sentences are not evaluated for truth in isolation, but rather with respect to a context. Evaluation of a sentence within such a context then results in the creation of a new context.¹³ With the shift of focus from a truth-based conception of content to an evaluation in terms of context change, the claim is that the denotational content of an expression can be defined in a unitary way as its context change potential, and that it is this shift that provides the needed flexibility in characterizing the intrinsic content of an expression. In several variants of this form of analysis, discourse-based anaphora is analysed in terms of a Russellian definite description with entailment of uniqueness (see Neale 1990, Heim 1990, Cooper 1996, van der Does 1996, Milward 1995, F. Breheny 1999), the anaphoric expression being defined as a context-update mechanism mapping some context containing a unique individual (hence the presupposed implication of existence) onto a new context, with the predication said to hold of that individual.

Such accounts however face the well-recognized problem that many examples fail to fulfil the condition of uniqueness while nevertheless constituting successful instances of pronominal construal. A pronoun (or definite NP) can be successfully used despite there being more than one individual in the context which could provide a possible value for the anaphoric expression, as in (1.29) said to refer to a coat which the speaker is just taking off in a changing room:

- (1.29) I like it, but it's too similar to this other one.

The response by Kadmon (1990) that the situation of coat liking is not the same as that in which the two coats are similar is unattractive, for the concept

¹²The only putative exception we know of occurs in Malagasy, which is argued by Zribi-Hertz and Mbolatianavalona (1999) to contain a pronoun restricted to bound variable interpretations.

¹³In Dynamic Predicate Logic and Stalnaker variants of interpretation, in which the context is a set of world-assignment pairs, the sentence is an update function which acts as a progressive filtering on such sets (see Stalnaker 1974, 1984, Heim 1982, Groenendijk and Stokhof 1991). However in Situation Semantics and DRT, with inference being defined over situations/discourse representations, the situation/discourse representation becomes part of the context, though DRT can also be interpreted in Stalnaker terms, as pointed out by Fernando (1994), van Eijck and Kamp (1997), and Kamp (1996).

of situation supporting the sentence appears to have been narrowed down to just the entities of which the relation is said to hold, and not to any broader partial model relative to which the assertion is evaluated.¹⁴

Also problematic are examples such as (1.30) said in a discourse situation in which there are two boys and two dogs, with just one boy holding one of the dogs too tightly:¹⁵

- (1.30) He's holding it too tightly.

How can a pronoun in such circumstances succeed in picking out the dog in question, given the single scenario? Given that there are two male individuals and two further non-human individuals, the assertion made is incorrectly predicted to be false. A move that is made in Situation Semantics, for example by Milward (1995), is to suggest that a situation must allow objects to carry indices, thus enabling the uniqueness relation between the linguistic expression and the semantic object which constitutes its referent to be retained. But this move involves treating individuals in situations as syntactic objects, discretely named by the index assigned to them, with a notable shift to a more representational perspective.

1.4.3 The Problem of Indirect Reference

Suppose we also set aside this problem, there remains yet the problem of indirect reference, but here the context may have to be broader than is provided by the previous sentence. (1.31) displays the problem:

- (1.31) John had a heart attack right outside the hospital, and they refused to treat him without an insurance card.

The problem is that, if the pronominal is taken to define an update from a context containing a fixed individual to a new context which incorporates the assertion made about that individual, there should be a clearly definable sense in which the first context contains that individual. But in this respect (1.31) is problematic. The hospital in the situation corresponding to the content of the first sentence of (1.31) is a building. Yet it is apparently sufficient to license the use of the pronoun *they*. Milward's (1995) proposal within situation-theoretic assumptions is to assume that situations relative to which sentences are evaluated contain not only the entities depicted, possibly carrying distinguishing

¹⁴Perhaps the most discussed case is that of the two bishops:
(i) If a bishop meets another bishop, he blesses him.

where to be true, both bishops must bless each other (see Kadmon 1990, Chierchia 1991, Heim 1990, Milward 1995).

¹⁵The example is discussed by Milward (1995) as a problem for the uniqueness implication of the definite NP; but the problem carries over directly to a pronominal variant, unless such a use is said to involve rigid reference to some individual demonstratively picked out – a move which side-steps the uniqueness problem only to re-confront the ambiguity problem. Such cases have been used to motivate a discrete concept of *resource situation* (see Cooper 1996). But, as with Milward's suggestion of indexed objects, the level of abstraction is such that the constructs posited are arguably nothing more than a representation of constraints to be met by some semantic object, rather than themselves constituting a semantic object.

indices, but also relations defined on those entities, the situation providing the interpretation for the sentence *John had a heart attack right outside the hospital* accordingly, having not only the entities John, a heart attack and a (unique) witness satisfying the expression *the hospital*, but also the relations between hospitals, doctors and medical conditions needing treatment, on the one hand, and between heart attacks and medical conditions needing treatment, on the other. But if such relations are defined as part of the situation satisfying *John had a heart attack right outside the hospital*, then the situation becomes quite unlike the very narrow concept of situation required to yield the right result for the earlier coat-choosing case.

Evidence that no extension of the notion of a situation to incorporate entities is sufficient for a complete characterization of anaphora is presented by cases in which the situation described may not include an individual, and yet reference to it by a pronoun can be successful. Consider (1.32) uttered to a friend who has just split up from his partner and is paranoid about meeting her at every turn:

- (1.32) I'm having a party. Don't worry: she won't be there.

The update made available by the first sentence (ignoring problems of modality) is some party. The only set of individuals that can in any sense be said to be specified as part of the context is some set as witness to the expression *a party*, or in some circumstances its complement set.¹⁶ But even allowing that the complement of a witness set might exceptionally form part of a specification of context, this would not alleviate the problem that this is not sufficient to guarantee the unique individuation of the person successfully referred to by the use of the pronoun *she* as part of the context relative to which it defines an update. What is required here is a concept of inference defined over the information provided by the utterance of the antecedent sentence – in this case along the lines ‘If A is having a party, C will be there. A is having a party. Therefore C will be there.’ But the need to construct premises from which the implication of the presence of C at the party is derived is driven by the presence of the pronoun itself. The specific entities depicted in the situation are in no sense available in the context within which the sentence is uttered. So the concept of context is quite unlike the content-independent context presumed in a semantics in terms of context-updates.¹⁷

¹⁶Note the use of the complement set as the basis for establishing anaphor-antecedent relations in negative sentences such as (i), an observation noted by Corblin (1996), and developed by Kibble (1997):

(i) They didn't invite a professor. She would have spoiled the party.

¹⁷This is the problem of presupposition accommodation (see Beaver 1997 for an overview). But notice that in this case accommodation must not have the effect of updating some party context to include the designated individual. Formal constructive theories of context have been proposed by van der Does (1996) and Rooth (1992, 1996).

1.4.4 Quantification

The final problem for the semantic account of anaphora is the use of quantified expressions which enable an antecedent to be made available for construal of a following pronoun that is not part of the interpretation of the quantified expression itself. In (1.33) two situations are described disjunctively, and the pronominal *her* then refers to some individual indicated in whichever of the two situations holds:

- (1.33) Either we can get a girl to read a poem, or we can get an older woman to rattle off a passage from Yeats. If it takes her 10 minutes,...

In such cases, one might argue, a pair of situations is introduced as a context of evaluation (as in Milward 1995), with no means of choosing between them.¹⁸ What is needed is a step of inference derivable from the two situations severally described to a third situation containing someone who reads for whoever ‘us’ is. But no such individual is available in the situation corresponding to the interpretation of the sentence. To side-step such problems, Milward suggests a concept of meta-level disjunction, without any ability to distinguish between the disjuncts. He posits an arbitrary name mechanism, the name in some sense picking out either one of the two individuals, both being ‘in attention’. But this is to directly incorporate proof-theoretic, and hence syntactic, constructs into the semantic characterization. It is simply the rule of Disjunction Elimination of natural deduction systems in a different guise (see Prawitz 1965, Lemmon 1965).

1.4.5 Syntactic Processes of Anaphora

Notwithstanding all these possible extensions to the semantic vocabulary, some cases of pronominal anaphora remain intransigent. Anaphora may be integrated into the syntactic system through the use of resumptive pronouns in both topic and focus constructions and in relative clauses, and all such phenomena are standardly taken to be unrelated to the central cases of anaphora. In certain circumstances, a pronoun may function as a gap, with the pronoun appearing to be subject to the very same subjacency constraints as syntactic gaps (Swedish – Maling and Zaenen 1981):

- (1.34) **Lisa talar jag med pojken som kysser (henne)* [Swedish]
 Lisa talked I with the boy who kissed (her)
 ‘As for Lisa, I talked with the boy who kissed her.’

¹⁸These cases are highly problematic for DRT also, which imposes the requirement of introducing a new discourse referent for each indefinite, thereby guaranteeing that there is not a uniquely determined antecedent for the pronominal. Moreover, the embeddability conditions for the conditional which follows require that all partial models in which the antecedent holds must be extendable to that of the consequent. But the individual to be constructed in that subsequent DRS cannot be identified as some girl or some woman as previously set up, because no such entity is introduced at the top level of the DRS constructed from the previous sentence. The analogous problem applies to Dynamic Predicate Logic, for which they are equally problematic.

In others, they serve as a means to enable gap devices to be available in those circumstances in which a syntactic gap is precluded (Shlonsky 1993, Aoun and Choueiri (1997), Aoun and Benmamoun 1998):

- (1.35) *qaraʔtu ll-maqālat llaðii saafara Š-Šaabu llaðii
read-I the article that travelled the young man that
kataba-ha [Classical Arabic]
wrote-it*

'I read the article that the young man who wrote it travelled.'

This type of phenomenon is not generally taken to be a problem given the syntax–semantics divide, but it should be noted that no semantically based explanation of anaphora purporting to provide a uniform explanation will provide an account for this type of anaphora, for the licensing of this form of pronominal construal appears to be determined by uncontroversially syntactic properties, and hence cannot be characterized in purely denotational terms. Yet this phenomenon runs right across different language families, suggesting a systematic interaction between certain types of syntactic processes and the general anaphora phenomenon, a puzzle which singularly lacks explanation if the two sets of phenomena are taken to be unrelated.

1.5 The Anaphora Solution – Towards a Representational Account

Though no one of these arguments taken on its own is a knock-down argument against characterizing an anaphoric expression in terms of its denotational properties – after all, ambiguity of linguistic expressions can always be invoked – the cumulative evidence leaves wide open the challenge of providing an account of anaphora which captures the phenomenon in an integrative way. In what follows, we shall be articulating an account of interpretation which is set within a representationalist methodology. We shall assume that linguistic input projects instructions for building up logical forms, and that these instructions, in combination with information independently established in context, yield some completed propositional structure. Any one such logical form is then combined with some larger set of premises, *the context*, to derive context-specific inferences (Sperber and Wilson 1995). This overall set of premises represents the full pragmatic content of that utterance (including so-called *contextual implications*), a set which then forms (part of) the context for the following utterance.¹⁹ On this view, the computational/lexical system of a natural language systematically underdetermines the set of possible interpretations for any sentence, and it is only interaction between pragmatic and system-internal actions which yields a context-particular representation of content as articulated in some logical form.

¹⁹Our focus in this book will be exclusively on the process of logical form construction, not on the deduction of contextual implications.

Against this general background, the diversity in pronominal interpretation can be modelled by articulating pronouns as projecting representations with only a skeletal indication of an interpretation – merely that of a place-holder to be substituted by some proper representation. The only structural restriction on this substitution process is one of locality; and antecedent representations respecting the specified locality restriction may be selected from a wide variety of sources – other terms in the same configuration (cf. (1.6), (1.7), (1.9) above), terms constructed during the processing of some previous utterance (cf. (1.8), (1.10)), terms representing objects in the utterance situation (1.30), or even terms derived by some context-specific inference (cf. (1.31), (1.32)). The manipulation of steps of inference to derive a term to take as antecedent (1.33) is not in principle problematic, given a deductive perspective on interpretation. Reflecting the supposedly syntactic nature of the interpretation task, there is no attempt, on this view, to provide a characterization of anaphora resolution in denotational terms, for semantic evaluation is defined not on the linguistic string itself, but on the logical form which results from the interpretation process of which anaphora resolution is a part.

All such statements about the interpretation task are of course nothing more than a promissory note without a formal characterization of both the system of representations in terms of which such interpretations are expressed and the process of building up such representations. Indeed, one of the purposes of this book is to articulate a formal model of the process of interpretation to give substance to some of these claims. It remains to be seen, when such a framework has been articulated, whether construals of anaphoric expressions which involve syntactic processes can be described in the same terms as what we might broadly refer to as discourse anaphora, and this question is taken up in chapters 4 and 6. However, whatever the answer to that question, the evidence of underspecification intrinsic to natural language expressions and the apparently structural nature of the enrichment process that resolves this underspecification suggest that the enterprise of setting out a model of natural language understanding in representationalist terms is well worth exploring in some detail. We now turn to the task of articulating such a model.