Language, thought and literal meaning

PhD thesis submitted to the
Department of English Literary and Linguistic Studies
University of Newcastle upon Tyne

Soke Wang Chng

NEWCASTLE UNIVERSITY LIBRARY
099 13089 5
Thesis L6507.

August 1999

Supervisor: Professor Noel Burton-Roberts

CONTENTS

Abstract v

Acknowledgements vi

Chapter 1 Literalness and the problem of "language use" 1

Chapter 2 Literalness and Chomsky's I-language 13

2 Introduction 13

2.1 Chomsky's internalist theory of language 15

2.1.1 I-language 18

2.1.2 Chomsky's internalist semantics 23

2.1.3 The "USE" of (I-)language 30

2.2 How exactly does the I-linguistic semantics bear upon the "literal use" of an I-language? 43

2.3 Chomsky on the "misuse of language" 50

2.4 Conclusion 57

Chapter 3 The distinction between linguistic and truth-theoretic semantics: literal meaning and relevance theory 61

3 Introduction 61

3.1 Relevance theory's Fodorean foundations 64

3.1.1 Relevance theory's departures from Fodor 68

3.1.2 Ostensive-inferential communication and the principle of relevance 77

3.2 Sperber and Wilson's account of verbal communication: logical form development 80

3.2.1 "Semantically incomplete" logical forms and the distinction between

linguistic and truth-theoretic semantics 90

3.3 Sperber and Wilson's account of verbal communication: propositional form identification 106

3.4 Conclusion 117

Chapter 4 The "misuse" of language 120

4 Introduction: phonetic vehicles and "linguistic" conventions 120

4.1 Carston's ad-hoc concepts: a pragmatic account of misuse and slips 123

4.2 Davidson's prior and passing theories 146

4.3 Dummett's reply to Davidson: language as social convention 155

4.4 Language use, literalness and convention 162

Chapter 5 A representational account of "literal meaning" and "language use" 177

5 Linguistic semantics without "linguistic" conventions: Burton-Roberts' representational conjecture 177

5.1 The representational conjecture 179

5.2 What is linguistic semantics the semantics of?: the nature and status of phonology within the representational framework 187

5.2.1 M-representation and the conventional vehicle-meaning relation 194

5.2.2 M-representations, Carruthers and conscious thought 209

5.3 A naturalist view of the relation between the linguistic and the "language of thought": the linguistic *as* "language of thought" 212

Chapter 6 The use of M-representations of linguistic expressions: some explanations, examples and questions 230

6 M-representation and the distinction between "semantics" and "pragmatics" 230

6.1 Conventional-(2) and unconventional M-representation of linguistic expressions 235

6.2 Non-iconic non-conventional M-representation of linguistic expressions 248

6.3 More on "linguistic codes" and "literal meaning" 256

6.4 The non-M-representation of linguistic expressions and the relation between the linguistic and the language of thought 262

6.4.1 M-representation and reference 270

Bibliography 278

Abstract

The notion of literalness in linguistics is based on the following assumptions:

- Linguistic expressions are vehicle-meaning pairs (since literal meaning has to be the meaning *of something*).
- Linguistic expressions have to be cognised and used (especially uttered) in order for their meanings to be *regarded as literally theirs*.
- "Linguistic" vehicle-meaning relations are fixed and autonomous -- rather than having particular meanings in virtue of being used to express those meanings, "linguistic" vehicle-meaning pairs are used to express certain meanings in virtue of having the meanings that they have.

This thesis criticises Chomsky's and Sperber and Wilson's attempts to establish the autonomy of "linguistic" vehicle-meaning pairs. I argue that

- Both Chomsky and Sperber and Wilson fail to distinguish "linguistic" semantics from the "real" semantics of what "linguistic" vehicle-meaning pairs are used to express.
- They persist in the idea that "linguistic" vehicles are specifically for being uttered (physically instantiated), thus defeating their own purpose of setting the linguistic absolutely apart from what it is used for.
- Neither Chomsky's internalist conception of language nor Sperber and Wilson's relevance framework is able to account for the phenomenon of "language misuse", i.e. the use of a "linguistic" vehicle to express the "wrong" meaning.

Burton-Roberts' representational conjecture is applied and developed in the presentation of an alternative non/ extra-linguistic account of "literal meaning" and "language use/ misuse". This account has it that neither "linguistic" vehicles nor "linguistic" vehicle-meaning relations are actually linguistic. It avoids the problems attending the notion of linguistic expressions as objects with sortally disjoint and arbitrarily conjoint properties (i.e. physically instantiable "vehicle" and mentally constituted "meaning"), and resolves the unease within Chomsky's Minimalist Program about the inclusion of phonology in I-language. Finally, by way of this resolution, I address some seemingly unrelated issues concerning vehicle-less "meanings" and the relations between language, thought and consciousness.

Acknowledgements

A Faculty of Arts studentship from my university and an Overseas Research Students award made it possible for me to work full-time on this thesis. My obligation to the Faculty of Arts of Newcastle University and the ORS Awards Scheme is hereby gratefully acknowledged. I am also grateful for the love and support of my family, and for the help and friendship which I have received from the staff of Newcastle University's linguistics department, and from my fellow postgraduates Zoe Boughton, Tina Fry and Irenie Rowley. The bulk of my thanks, however, go to Noel Burton-Roberts: an infinitely patient supervisor and an incomparable teacher. It is inconceivable that anyone other than he should be the dedicatee of this thesis, if only because of what I owe him in ideas and inspiration. Hopefully his views have not been misrepresented in the following pages.

1

Literalness and the problem of "language use"

This thesis was originally intended as an investigation of the phenomenon of *metaphor*. However, it soon became clear to me that folk concepts as well as theoretical definitions of metaphor rest upon assumptions about things which are equally (if not more) difficult to pin down, and in need of elucidation: *language* and *meaning*, *thought* and *consciousness*, and, most particularly, the notion of *literalness*. It is my belief that literalness requires as much discussion as metaphor, if only because the concept of metaphor only makes sense, or has any substance at all, in relation to that of literalness. Certainly it seems very difficult to talk about metaphor without making any reference to literalness, and (possibly) vice versa. It is as if one was inextricably linked with, because conceived in terms of, the other.

This is true even of those theories which Ortony (1993: 2) calls "constructivist" -- i.e. accounts of metaphor based on the hypothesis that "meaning has to be constructed rather than directly perceived". Ortony (ibid) maintains that "constructivist" theorists like Lakoff (1987, 1979/93, Lakoff and Johnson 1980), in claiming that metaphor is an essential feature of language and thought, tend to undermine the distinction between the metaphorical and the literal. Trivially, any account or discussion of metaphor might include some notion of literalness, if only to deny or reject it, or to proclaim it as inferior to metaphor as a means of communication or expression. For example, the "constructivist" might describe an apparently unworkable and impossible "literal language" which precisely and unambiguously depicts the world, as a way of illustrating what he considers to be the contrasting (and contrastingly plausible and pervasive) nature of metaphor.

In my view, the notion of literalness has another more fundamental and substantial role in "constructivist" accounts of metaphor. Lakoff is a case in point. Briefly, his theory is that metaphor is a mechanism that plays a very central role in the workings of

the mind: it maps concepts on to other concepts, and in doing so enables us to grasp yet more concepts. But, contrary to Ortony's assertion, Lakoff's account of metaphor *does* require a separate notion of literalness. According to Lakoff, it is from "literal concepts" (1979/93: 205) or "emergent concepts" (Lakoff and Johnson 1980: 60) that the mapping process begins and builds in complexity (and, presumably, in non-literalness). There is no sense in which the concepts which arise from the metaphorical mapping mechanism are less "normal" or important than those that do not, or vice versa. But there is no doubt that Lakoff's notion of literalness is *pre*-metaphorical. As he puts it: "Metaphorical understanding is grounded in non-metaphorical understanding." (1979/93: 244).[1]

Similarly, I.A. Richards argues that metaphor is "the omnipresent principle" of language's "normal mode of working" (1936: 90), but nevertheless distinguishes between the two conceptual "halves" of a metaphor: the *tenor*, the "underlying idea" which the other idea, the *vehicle*, can be used to "mean", or "describe or qualify" (ibid: 99). This unavoidably implies that the metaphorical combining and reacting of vehicle with tenor, while assumed to be the "normal" way in which language works, is made possible only by the presence of literal or at least pre-metaphorical meanings, expressed in non-metaphorical ways, which function as the vehicles. For instance, Richards' analysis of 'crawling' in Hamlet's utterance "What should such fellows as I do crawling between earth and heaven?" as a metaphorically used vehicle presupposes that 'crawl' does have an original, literal meaning which constitutes the vehicle:

> Hamlet... may *crawl literally* -- as babies and big-game hunters undoubtedly do at times -- but... there is an unmistakeable reference to other things that crawl, to the motions of foul insects, to vermin, and this reference is the vehicle as Hamlet... [is] the tenor. (ibid: 119, my italics)

---

[1] Actually it is not clear why Ortony cites Lakoff 1979/93 as a "constructivist" account, or at least an entirely constructivist account, since Lakoff does have a very clear distinction between concepts which are "constructed" by the metaphorical mapping process, and "literal concepts" which are *not*.

Thus I argue that there is a paradoxical side to theories which hold that metaphor is the norm in language and thought, and that language and thought are essentially metaphorical in nature. If the difference between metaphor and literalness is a matter of degree rather than of quality, then *any* meaning, and perhaps the means of expressing any meaning, ought to be ever shifting and provisional (though probably far from arbitrary). In other words, there should be no distinction between literal meaning and meaning which arises from metaphor, because the assumption would be that ALL meaning arises from some ongoing metaphorical process.

But is this assumption consistent or compatible with what is inarguably the main purpose of any theory of metaphor -- to define what metaphor is? Whereas Ortony claims that the "constructivist" approach tends to blur the literal-metaphorical distinction, my view is that such a distinction is created or presupposed by any attempt to circumscribe the phenomenon or process of metaphor, or even to single out specific "examples" of metaphor. It seems that the mere existence of metaphor cannot be affirmed unless one invokes (directly or indirectly) some concept of literalness, or at least the idea that there are meanings other than those which arise from metaphor itself. Take Richards, who assumes that 'crawl' has a literal meaning to do with the motions of babies and insects, and whose test for metaphor is that we should be able to detect at least two "ideas" when presented with a metaphorical expression (i.e. that we should be able to distinguish tenor from original vehicle). Without the literal meaning to serve as the vehicle, there would be nothing for the tenor (the idea of Hamlet and his actions and concerns) to interact with, no "borrowing between and intercourse of thoughts" (Richards 1936: 94) -- and no resulting metaphorical meaning. In short, without the concept of literal meaning, it would be impossible for Richards to say how he recognised the metaphorical nature of 'crawl' in the *Hamlet* quote, let alone explain how it came about.

To reiterate and expand on what may appear to be a rather trivial and obvious point: if a theory proceeds from the premise that metaphor has to do with meaning and/ or the expressing or constructing of meaning, then that theory must include some

notion of literalness or literal meaning. And not just any notion of literalness -- needless to say, there is more to the literal than that which is separate from and prior to the metaphorical.

To begin with, the notion of literalness obviously presupposes the notion of certain objects or units -- specifically E(xternally)-physical or E-instantiable objects/ units[2] -- *having meaning* or *being the vehicles of* meaning. Generally, these objects or units are thought of as *expressions*, in the sense of being *in themselves* expressive of certain meanings. It is the meaning *of a particular vehicle or expression* which is assumed to be *literal*. Metaphorical meaning, on the other hand, is never unequivocally described as the meaning of anything (except "metaphors", of course, whatever metaphors are). Perhaps metaphorical meaning might be conveyed or expressed by, say, the uttering of an utterance. But, when it comes to *the meaning of that utterance*, it is more often the case that only the literal would be taken into account.

Vehicles of meaning tend to be regarded as being USED (by agents) to express certain meanings, but the important point is that the vehicles *in themselves* -- as vehicles of meanings -- are supposed to express whatever meanings they carry, rather than what they are used to express. Thus the notion of literalness also involves a pretheoretical opposition to Humpty-Dumpty's insistence that, when he uses a word, it means what he chooses it to mean. However, it has to be pointed out that meaning M is regarded as the literal meaning of vehicle U, not only in virtue of being the meaning of U, but also in virtue of being recognised or regarded as the meaning of U. So, on the one hand, the literalness of meaning M of vehicle U presupposes a vehicle-meaning relation that is independent of what humans think or do. But, on the other hand, M can be said to be literal only in virtue of the U-M relation's place or role in human thought and activity.

---

[2] I am following Chomsky in my use of "E" to stand for "external". See the following chapter on Chomsky and his problematic (1986) concept of "E-language".

It is LANGUAGE which is usually assumed to be the primary, and possibly most natural,[3] vehicle of meaning. In fact the term "language" has become almost synonymous with "vehicle of meaning", for "language" not only has its more specific (including the strictly theoretical) definitions, by which only certain entities are regarded as linguistic, but is also used to refer to anything which is considered to be expressive of meaning (e.g. "a musical language", "a language of cinema"). Furthermore, the vehicle-meaning pairs which supposedly constitute *linguistic expressions* are generally assumed to be determined by sets of principles which are somehow fixed -- which pre-exist, which do not need to be constructed, which humans simply *have*. I have argued that assumptions of and about literalness are integral to theories of metaphor, and therefore tend to be made in respect of how metaphor is conceived. But while meaning is considered metaphorical in origin only if it is arrived at through the interaction of literal meaning with metaphorical process, it is common for literal meaning to be considered independently of metaphor, and in relation to *language* instead. My point, which may again seem trivial and obvious, is this. Meaning which is said to be literal is in most cases associated with linguistic meaning, in virtue of the latter's being paired with particular phonetic vehicles. In other words, it is linguistic meaning that is most often associated with the property of literalness, mainly because linguistic meaning seems to be characteristically and undeniably the meaning OF certain vehicles.

As I pointed out above, however, literalness involves more than vehicle-meaning pairs -- it also involves an individual's awareness and application of those vehicle-meaning relations in the process of expressing or accessing meaning. In short, literalness (as well as non-literalness) also involves the USE of the vehicle-meaning pairs which linguistic expressions are generally assumed to be.

This thesis is mainly concerned with the problems surrounding:

---

[3] I am using "natural" here in the vague and gradable sense of *not exceptional*, rather than the non-gradable and more specific sense which Chomsky (e.g. in Chomsky 1995a) uses to indicate that his I-language is a part of the natural world, and an object of naturalistic inquiry.

(1) the autonomous vehicle-meaning pairs which linguistic expressions are assumed to be,

(2) the use of such vehicle-meaning pairs, and

(3) the two loci of meaning which follows from the distinction between (1) and (2).

Note that there would be two loci of meaning even if we assumed that linguistic expressions are used to convey only the meaning which they possess in themselves. Take a speaker who uses the linguistic expression, L, to convey propositional meaning M. Even if L was used to convey M in virtue of L having a semantics which consists of propositional M, the speaker could still have entertained M independently of L *before* picking out L to use to convey M.

In fact, in both of the theoretical frameworks within which I consider the notion of literalness, an absolute distinction is made between the meaning of L and propositional M expressed through the use of L. The first is Chomsky's internalist theory of language (chapter 2), according to which the **linguistic semantics** of **linguistic expressions** generated by **I-language** is distinct from the **propositional meaning** which the **performance systems** are concerned with. The second is Sperber and Wilson's relevance theory (chapter 3), which distinguishes between the **linguistic semantics** of the **linguistic codes** over which **language modules** compute, and the **truth-theoretic (or "real") semantics** which the **central processes** deal in. In both cases, the propositional meaning or truth-theoretic semantics appears to be regarded as the semantics, not of language as such, but of some ***language of thought*** (to use Fodor's expression). Its relation to language is that it may be expressed or cognised through the USE of some linguistic expression, as effected by the performance systems or central processes.

However, while making such *distinctions* between language and thought, neither Chomsky nor Sperber and Wilson manage to provide an independent account of how

the two are *related*. The notion of literalness is central to my argument. Whereas the phenomenon of metaphor highlights the distinction between language and language use, the concept of literalness -- based as it is on the notion of *vehicles being used in virtue of what they mean* -- highlights the problem of how we might bridge the empirical and conceptual gap between autonomous linguistic expressions and the mental processes which effect the use of linguistic expressions. This problem may be formulated in terms of the distinction/ relation between linguistic semantics and what the relevance theorists refer to as truth-theoretic or "real" semantics: how exactly is the linguistic semantics of some linguistic expression *used* to express a particular "real" propositional meaning which may be regarded, if not as "the literal meaning of that linguistic expression", then at least as "the meaning expressed by the literal use of the linguistic expression"?

The idea that a linguistic expression may be used literally -- to express a meaning which is, if not identical to, then at least constrained by, its linguistic semantic properties -- also raises questions about the nature of a linguistic semantics which is closely related to, yet distinct from, the propositional meaning expressed through its literal use. It seems to me that both Chomsky and the relevance theorists conceive of linguistic semantics in terms of what it is used for, and in doing so defeat their original purpose of distinguishing linguistic semantics from truth-theoretic propositional meaning. Although it is usually the case that an object is used for a particular purpose in virtue of inherently possessing particular properties, it is also true that certain characteristics tend to be attributed to that object as a result of its being used. For a start, the object would be thought of as "usable for that particular purpose". Chomsky seems to be carefully separating language from what language is used for, when he remarks:

> In general, it is not the case that language is readily usable or "designed for use." The subparts that are used are usable, trivially... (1995b: 18)

Sperber and Wilson, on the other hand, claim that language[4] is specifically *for* verbal communication (1986a/ 95: 173-4). In chapters 2 and 3 I argue that neither Chomsky nor Sperber and Wilson manage to avoid the implication that, instead of being used to express propositional M in virtue of having a particular semantics, a "linguistic expression" has a particular semantics -- i.e. is regarded as an *expression* -- in virtue of being used to express M. This confusion of "linguistic (semantic) properties" with "what the linguistic is used to express" is compounded by the suggestion, made by Chomsky (e.g. in 1992a: 115-6) as well as Carston (e.g. in 1988: 177, 1998: 64-5), that the mental resources involved in the use of linguistic expressions to express truth-theoretic meaning also constitute the basis for (or the source of) the semantics of those linguistic expressions.

Rather than contrasting the literal use of language with the metaphorical use, I discuss in chapter 4 the apparently straightforward but in fact highly problematic concept of the so-called "*misuse*" of language. The concept of language misuse appears to follow from the notions of language and language use described above. However, rather than developing those notions, it reveals some of their fundamental deficiencies and contradictions. As it is generally understood (e.g. in Chomsky 1992a and 1995a, and in Carston 1998), the concept of language misuse presupposes that there is a particular use of a particular linguistic expression which is the correct use, and which is the correct use in virtue of being in some way constrained by the *semantic* properties of the linguistic expression. For example, the misuse of the linguistic expression 'flaunt' in an utterance

(4) She always flaunts the rules.

---

[4] Note that by "language" here I mean the "external languages" which constitute relevance-theoretic "linguistic codes". Sperber and Wilson distinguish such "external languages" from the "internal languages" which are "an essential tool for the processing and memorising of information" (1986a/ 95: 173).

is usually taken to consist in the (mis)use of the E-physical (phonetic) vehicle of the expression to express some truth-theoretic propositional meaning (i.e. the concept *flout*) which is at odds with the linguistic semantics carried by the vehicle. On the one hand, this presupposes that there is a truth-theoretic meaning, namely *flaunt*, which is consistent with the semantic properties of linguistic expression 'flaunt', and expressed through the (correct, literal) use of that linguistic expression. On the other hand, note that, in failing to (intend to) express the truth-theoretic meaning *flaunt*, the speaker has actually failed to make use of the linguistic semantics of the linguistic expression 'flaunt'. Nevertheless, his production of the phonetic vehicle of the linguistic expression is considered (at least by Chomsky (1992a, 1995a) and Carston (1998)) as sufficient (and obviously necessary) for the linguistic expression 'flaunt' to have been (mis)used.

What this underlines, in my view, is that the misuse -- and therefore the use -- of a linguistic expression is generally regarded as consisting primarily in the production of its E-physical vehicle rather than the deploying of its semantic properties. This notion of language use/ misuse seems to be inextricably bound up with the view of linguistic expressions as having E-instantiable *vehicles*, in virtue of which linguistic expressions are to all intents and purposes inherently E-produceable for the purpose of conveying/ carrying meaning. Furthermore, it is unable to accommodate cases of so-called "misuse" in which the person who utters, say, (4) not only intended to express the concept *flout*, but intended to do so via the specific and (for him) *literal* use of [flɔ:nt], regardless of the "linguistic" convention which states that the phonetic form is the vehicle of the semantics which corresponds to the concept *flaunt*.

At this point I feel it is only fair to warn the reader that my discussion and criticism of (1)-(3) is informed by Burton-Roberts' 1994 **representational conjecture** (outlined in chapter 5), which in my view provides a far more coherent account of so-called "linguistic" vehicle-meaning pairs and their use. The conjecture is that E-physical utterance phenomena are *non-natural and intentional representations* of linguistic expressions, in (much but not exactly) the same way that a drawing of a square in a map is a representation of a building. Burton-Roberts stresses that the "representation"

of the representational conjecture (henceforth RC) is *not* the "representation" of Chomsky's (1995a, 1995b) "linguistic levels of *representation*". "Representation" in Chomsky's sense is not relational but *constitutive*. For example, a "phonological representation" of a Chomskian linguistic expression simply *is* the phonology of the linguistic expression. For the RC, it is crucial that a representation "is distinct from, and is not constitutive of, what is represented" (Burton-Roberts 1994: 192). BR distinguishes his "representation" from Chomsky's by labelling it with an "M" -- hence *M-representation*. "M" stands for Magritte, alluding to Magritte's *La Trahison des Images* and its reminder that *Ceci n'est pas une pipe* -- i.e. that the painting contains only a representation of a pipe, and not the pipe itself.

Thus BR posits that E-physical utterance phenomena are produced by speakers, writers, etc in aid of E-physically *representing* linguistic expressions. This means that, on the representational conjecture, the *E-phonetic* is clearly not linguistic. However, it has to be stressed that the arguments of chapters 5 and 6 follow on from my claim that it is not just the E-phonetic, but *phonology* as well, which is M-representational and therefore not linguistic. There are accounts and applications of the RC which address specific issues in syntax, semantics and pragmatics -- e.g. linear precedence (BR 1999a), so-called "semantic ambiguity" (BR 1994, 1999b), quotational mention (BR 1993, 1999b), the Grelling paradox (BR 1999b) and utterance accent (Chapman 1998). All these (especially the earlier BR 1994) seem to take it for granted that the phonological is linguistic, radically internalist, and conventionally M-represented by the (non-linguistic) E-phonetic. But recent work within the representational framework (BR 1998, BR and Carr 1999, BR 2000, Carr 2000) has focused specifically on the nature and status of phonology. Burton-Roberts and Carr (1999: 399) remark: "The conjecture does not, of itself, deliver a decision on [the nature and status of the phonological]." Nevertheless, they point out that the tensions in phonological theory are connected to the longstanding dispute over whether the linguistic is a natural, innate state of mind/ brain (*Language*), or constituted by socioculturally-determined systems of conventions that are specifically for (E-behavioural) communication

("*languages*", e.g. English, Italian, Swahili). BR and Carr argue that it is only "languages", and not Language, which can have a phonology. In fact, phonology is necessary for the E-instantiation or E-realisation of "languages" in communication, while the characteristic Saussurean arbitrariness and convention of these "languages" obviously arise from the way the phonological is paired with the semantic.

The relevance of the RC to this *Language vs. "languages"* issue is that it provides an account of how one is related to the other -- and, more specifically, of how the phonological belongs exclusively with "languages", but is also relevant in respect of Language. BR and Carr (BR and Carr 1999: 399-404, BR 2000) argue that "languages" are M-representational of Language, and that phonology is a (non-linguistic) internalised attitude to the (non-linguistic, M-representational) E-phonetic phenomena which constitute speech. In short, the phonological constitutes those "languages" which are regarded as "spoken languages" (e.g. English, Italian, Swahili).

Applying all of the above to (1)-(3), I claim that vehicles and (Saussureanly arbitrary and conventional) vehicle-meaning pairs are not linguistic, but *M-representational* of the linguistic. I also claim that the various "languages" with their various sets of vehicle-meaning pairs, are in fact systems of representational conventions. Such systems are neither linguistic nor intrinsically fixed or autonomous. Instead, they arise from the volitional use, *not of linguistic expressions*, but of E-physical objects to represent linguistic expressions. Like any other volitional human activity, this use of the E-physical to represent the linguistic is subject to external (especially sociocultural) factors.

But whereas much of BR and Carr's latest RC-based work is concerned with phonology -- i.e. with the "*vehicle*" of the vehicle-meaning pair -- this thesis takes their claims about the non-linguistic (M-representational) nature of phonology as a starting point for projecting the RC into other, more general areas of inquiry that have to do with the "*meaning*" of the vehicle-meaning relation. I argue that it is the *meanings*, left behind by the vehicles' shift to non-linguistic M-representational status, which constitute the I-linguistic (or Language). Note that these meanings are not the

meanings/ semantics *of anything*. In my opinion, the logical conclusion of the RC is this: linguistic expressions are vehicle-less syntactico-semantic objects, indistinguishable from the "real" semantics of the language of thought. Such a view of the linguistic has implications for the linguistic's relation with thought and consciousness, and for the architecture of mind in general, which may not be considered acceptable even by others who subscribe to the representational conjecture. These implications will also be discussed in chapters 5 and 6.

To conclude, a word about the notational conventions of this thesis. I begin by using single quotes for the "linguistic expressions" of Chomsky and relevance theory, double quotes for utterances of these "linguistic expressions", and italics for the semantics of "linguistic expressions" as well as for the "real" semantics of what is conveyed via particular uses/ utterances of "linguistic expressions". For example, when Bill produces the utterance "thing", he is using (the "linguistic expression") 'thing' to convey *thing*. Note, however, that "linguistic expressions" which can be uttered are utterable/ usable/ E-instantiable in virtue of having vehicles, i.e. phonological properties. As mentioned above, one of the main arguments of chapter 5 is precisely that vehicles and vehicle-meaning pairs are M-representational and therefore not linguistic. It follows from this that the utterable "linguistic expressions" indicated by single quotes are not of natural, innate, I-cognitive *Language*, but of E-instantiable, M-representational *"languages"*. The development of this argument will be reflected by the changes in my notational conventions, starting from section 5.3. Instead of single and double quotes, there will only be braces and italics -- braces for M-representations (e.g. {thing}), and italics for the vehicle-less, phonology-free syntactico-semantic objects which constitute linguistic expressions (e.g. *thing*).

2

Literalness and Chomsky's I-language

2 Introduction

This chapter is mainly concerned with a certain view of the relation between literalness and language, the assumptions underlying such a view, and the question of whether these assumptions -- about language as well as literalness -- are coherent, both in themselves and as the basis for associating literal meaning with linguistic semantics. My approach is to consider the nature of literalness within the framework of a particular theory of language. I have chosen Chomsky's, for the reason that it appears to be the most detailed and uncompromising view of language as an autonomous, unchanging set of principles which exists and operates with absolutely no reference to anything other than itself and its own inherent properties. According to this view, language is quite separate from the mental heterogeneity which characterises metaphor and non-literalness in general, and therefore an apparently ideal basis for a concept of literal meaning that is truly distinct from metaphorical meaning.

However, I also argue that there may be a fundamental problem with Chomsky's theory of language (discussed in 2.1) -- a problem which is highlighted by my attempt to apply the theory to the notion of literalness. Chomsky could be said to have reconstructed the pre-theoretical autonomy of vehicle-meaning pairs as purely internalist linguistic expressions (LEs) generated by I-languages which are supposed to be entirely innocent of human activities and concerns -- and, more specifically, entirely innocent of *how I-languages are USED, or what they are used for*. I argue that, instead of confirming and strengthening the autonomy of the linguistic "vehicle-meaning pairs", Chomsky's

internalist conception of language seems to lead to an even more complex entanglement of linguistic properties with aspects of language use.

To begin with, the austerity of Chomsky's I-language is such that the speaker-hearer who has an I-language is (according to Chomsky) able to consciously access it *only by means of using it.* (See, for example, Chomsky 1995b: 17.) It follows from this that speaker-hearers can only conceive of LEs in terms of how they are used. In other words, Chomskian I-language is *not "designed for use"* (Chomsky 1995b: 18, my italics), but can only be consciously accessed *by being used.* Furthermore, the austere and internalist semantics of Chomsky's I-language does not in itself consist of the sort of propositional meaning which may be regarded as literal, and has to be *used* to express that sort of propositional meaning. All this, I argue in 2.2, has the effect of subordinating I-language, and I-linguistic semantics in particular, to the speaker-hearer's use of I-language. I also argue that Chomsky himself deviates from his internalist perspective by defining and describing I-linguistic semantic properties only in terms of the propositional meanings which they are conventionally and extra-linguistically *used* to express. This comes across particularly clearly in Chomsky's discussion of the "misuse of language", as I attempt to show in 2.3.

It must be stressed that Chomsky himself almost never addresses the issue of literalness, or at least not directly. Indeed the entire Chomskian framework is very retiring on the question of linguistic semantics and propositional meaning in general -- perhaps because of the problems mentioned above. This means that any attempt to connect Chomskian linguistics with the notion of literalness is generally my own, unless I indicate otherwise.

2.1 Chomsky's internalist theory of language

Although Chomsky's theory of language has developed and changed not a little in the past forty years or so, it is nevertheless possible to pick out certain guiding ideas upon which his views have been consistently based. The most fundamental of these -- and I think the vaguest -- is that there is a discrete component of the human mind/ brain, the LANGUAGE FACULTY, dedicated to language (and that language therefore belongs in the natural world). As Chomsky himself points out in his recent *Minimalist Program*, this is one of the "underlying factual assumptions" (1995b: 2) running through his work since the 1950's.

But Chomsky's (A) positing of a specific, innate language faculty which "provides an array of capacities that enter into the use and understanding of language" (ibid: 167) is, in my opinion, not as important as (B) his view of LANGUAGE itself as being mentally constituted. It is (B) which makes Chomsky's theory of language an explicitly INTERNALIST theory, one that is diametrically opposed to externalist theories like the behaviourists'.

The externalist view holds that language consists only of the physical: physical utterance phenomena are all there is to language, and sentences are nothing more than potential utterances or theoretical constructs, abstractions from utterances actually produced and interpreted by speaker-hearers. Thus language in Bloomfield 1926, for example, is "the totality of utterances that can be made in a speech community". This is *not* inconsistent with (A), since the "capacities that enter into the use and understanding of language" may be construed as the ability to carry out certain inductive processes on certain categories of physical phenomena (i.e. utterances). But with (B) Chomsky is claiming that language is *not the thing which humans have the capacity to produce and interpret, but the (mentally constituted) capacity itself.* For Chomsky, language is COMPETENCE, "the speaker-hearer's knowledge of his language" (1965: 4), not

PERFORMANCE, the "actual use of language in concrete situations" (ibid) or the "practical ability to speak and understand" (1986: 9).[1] In *Minimalist Program* (1995b: 14), competence is "knowledge and understanding", while performance is "what [the speaker-hearer] does with that knowledge and understanding" -- language being, of course, competence rather than what is done with competence (i.e. performance).

Chomsky also distinguishes between what he calls I-LANGUAGE and E-LANGUAGE. This does not exactly correspond with the competence-performance distinction -- I-language appears to be equivalent to competence, but E-language is more accurately defined as the E-physical products of performance. I.e. Bloomfield's "totality of utterances that can be made in a speech community" is in fact Chomsky's E-language. The important point is that, for Chomsky, language is *not E-language* --

> The E-language that was the object of study in most of traditional or structuralist grammar or behavioral psychology is now regarded as an epiphenomenon at best. (1986: 25)

-- but I-language --

Both make this pt. too - refs?

[1] Here is a foretaste of problems to come. Chomsky's internalist concept of language as competence rather than performance is in fact at odds with his own definition of competence as the speaker-hearer's *knowledge of language* -- simply because language cannot simultaneously be the speaker-hearer's knowledge of L, and the L of which he has knowledge.(This issue is also discussed in George 1989.) My point also applies to the quote in (A) about the language faculty providing "an array of capacities that enter into the use and understanding of language" (1995b: 167). According to Chomsky's competence-performance distinction, language IS the array of capacities, and therefore cannot at the same time be what is "used and understood" -- unless, of course, the capacities provided by the so-called language faculty are not *linguistic* capacities. Even then, language as defined by Chomsky is not something which a speaker-hearer *knows* or *understands* -- as he himself points out, "Note that Jones has this knowledge [of language] whether or not he is aware of these facts about himself; it may take some effort to elicit such awareness, and it might even be beyond Jones's capacities" (1995b: 17). I suspect that the "language" which Chomsky describes the speaker-hearer as "understanding" and having "knowledge" of is in fact Bloomfield's "totality of utterances"; and that this use of the word "language" is just one of the many ways in which the externalist notion haunts Chomsky's supposedly internalist theory. See section 2.2.

> The central task is to find the basic elements of I-language -- henceforth, language. (ibid: 51)

E-language --

> ... where *E* is to suggest "external" and "extensional"... (1995b:16)

-- consists of utterances, the products of human behaviour, and therefore involves or is determined by

> ... complex and obscure sociopolitical, historical, cultural, and normative-teleological elements... which plainly lie far beyond any useful inquiry into the nature of language... (1992a: 102)[2]

The term "I-language", on the other hand, was

> ... chosen to indicate that this conception of language is internal, individual, and intensional... (1992b:221)

Chomsky describes I-language as "a state of mind, realised in some arrangement of physical mechanisms [of the brain]" (1986: 40), and as "[having] no objective existence apart from its mental representation" (1972: 169).

In short, Chomsky's main claim is that his internalist theory represents a shift

> ... from behavior or the products of behavior to states of the mind/ brain that enter into behavior. (1986: 3)

---

[2] Burton-Roberts (1994: 196) points out: "If so-called E-language really is of no significance to the theory of language, then the term "E-*language*" seems profoundly inappropriate -- as does Chomsky's continued reference to "linguistic behaviour"..."

-- "behavior" and its "products" being, respectively, what Chomsky calls performance and E-language, and "states of the mind/ brain" constituting I-language.

2.1.1 I-language

In this section I discuss I-language in greater detail, with particular attention to Chomsky's theory of I-linguistic semantics and how it is affected by the relation between I-language and what Chomsky calls the "performance systems".

Firstly, it is important to note that, while it is certainly not a set of utterances, neither is an I-language a set of sentences or linguistic expressions (LEs). Nor does it *manufacture* LEs, or bring them into existence in some active or mechanical way, even though Chomsky describes it as *generating* LEs. This is because Chomsky uses the term "generative" to mean something like "explicit":

> If the grammar [i.e. the I-language] is... perfectly explicit -- in other words, if it does not rely on the intelligence of the understanding reader [or speaker-hearer] but rather provides an explicit analysis of his contribution -- we may... call it a *generative grammar*. (1965: 4)
>
> ... "generative" means nothing more than "explicit"... (1986: 3)

The "generating", "computing", "mapping" etc performed by an I-language is in no way sequential or temporal. Perhaps the I-language should not even be referred to as "performing" anything (as in "performing a procedure", for example). To be entirely true to Chomsky's definition of "generative/ generate", an I-language cannot be described as assembling its LEs step by step from their constituent parts. Rather, it merely *specifies* what those LEs are or consist of. Another way of putting it is that I-language is nothing more than "a *description* of the ideal speaker-hearer's intrinsic competence" (1965: 4, my

italics). Chomsky's very specific use of the terms "generative" and "generate" has to be borne in mind when we come across verbs like "compute", "map", "form", "construct", etc in his accounts of I-language.

Thus the LEs generated by an I-language are not the *constituents* of that I-language, i.e. an I-language is not a set of LEs. And since they exist in virtue of being defined (rather than constucted) by the I-language, nor are they, strictly speaking, the I-language's *products*.

According to Chomsky's Minimalist Program (1995b), the I-language consists of (i) a COMPUTATIONAL COMPONENT (C_{HL}) and (ii) a LEXICON. C_{HL} determines computations or DERIVATIONS which form STRUCTURAL DESCRIPTIONS (SDs), while the lexicon characterises the lexical items which appear in the SDs. (1995b: 20). Although Chomsky occasionally makes it sound as if SDs and LEs were two separate entities --

> When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. (1965: 9)
>
> ... the conception of a language as a generative procedure that assigns structural descriptions to linguistic expressions... (1992a: 103)
>
> Each SD... specifies the full array of phonetic, semantic, and syntactic properties of a particular linguistic expression. (1995b: 15)[3]

[3] A rather more disturbing quote concerning the relation between LEs and SDs is

> Jones has mastered a generative procedure that associates with utterances structural descriptions, including semantic properties, and has other capacities of mind that allow him to *produce and interpret linguistic expressions making use of these structural descriptions*. (1995a: 121, my italics)

-- in which Chomsky appears to be using the phrase "linguistic expressions" to refer to E-physical utterance phenomena.

-- the fact is that SDs *are* what I-languages generate, i.e. LEs:

> We may take the linguistic expressions of a given I-language to be the SDs generated by it. (1992b: 221)

> One component of the language faculty is a generative procedure (an I-language, henceforth language) that generates structural descriptions (SDs)... These SDs are the expressions of the language. (1995b: 167)

By Chomsky's definition, an SD consists of a pair of representations or symbolic objects -- these are π, the PF ("phonetic form") representation, and λ, the LF ("logical form") representation:[4]

> The elements of these symbolic objects can be called "phonetic" and "semantic" features, respectively, but we should bear in mind that all of this is pure syntax, completely internalist... (1995a: 19)

π and λ, which together constitute an SD, are formed at PF and LF respectively. The two important characteristics of PF and LF are that (i) they are levels of representation or symbolic systems to π and λ's representations or symbolic objects; and (ii) they are INTERFACE levels which interface with particular PERFORMANCE SYSTEMS. As their name indicates, *performance* systems are to do with performance rather than competence -- i.e. what the speaker-hearer does with his competence/ I-language, or his *use* of his I-language. In Chomsky's words:

> ... performance systems... make use of [information stored in the I-language] for articulation, perception, talking about the world, asking questions, telling jokes, and so on. (1995a: 12)[5]

---

[4] Chomsky states that "representation" here is "not to be understood relationally, as "representation of"..." (1995a: 53). Within his internalist framework, representations like λ and π are "postulated mental entities" (ibid.), and do not *represent anything*.

[5] The notion of I-language being *made use of by performance systems* has to be considered in the light of an earlier remark in the same paper (Chomsky 1995a: 8):

In short, performance systems have to do with *behaviour* (i.e. performance) and the *products of behaviour* (i.e. E-language). PF interfaces with the ARTICULATORY-PERCEPTUAL (A-P) performance systems, and LF with the CONCEPTUAL-INTENTIONAL (C-I) performance systems.

Together, these performance systems and the I-language form what Chomsky calls the LANGUAGE FACULTY.[6] Whereas the I-language is "a cognitive system that stores information", the performance systems "access that information and use it in various ways" (1995b: 2).

So how does an I-language generate an SD? The following account is taken from one of the more recent versions of Chomsky's linguistic theory, his Minimalist Program (1995b).[7] An operation, presumably applying to the lexicon, forms a NUMERATION N:

> ... a set of pairs (LI, *i*) where LI is an item of the lexicon and *i* is its index, understood to be the number of times that LI is selected. (1995b: 225)

> More cautiously, we may say that in appropriate circumstances *people* think, not their brains, which do not, though their brains provide the mechanisms of thought. I may do long division by a procedure I learned in school, but my brain doesn't do long division even if it carries out the procedure.

Assuming that the performance systems are as much a component of the human mind/ brain as the I-language, this implies that ultimately it is the person -- and not his performance systems -- that makes use of his I-language.

[6] Note that Chomsky is not consistent as to whether or not the performance systems are included in the language faculty. The "language faculty" of Chomsky 1975(: 32-4), for example, consists only of principles specifying linguistic form. Burton-Roberts and Carr point out (1999: 387; the references cited are consistent with my bibliography) that "Chomsky sometimes has it that the linguistic includes... a "performance system" (e.g., 1995a: 12, 1995b: 2), at others that it is "embedded within" such a system (e.g., 1995b: 15, 158)."

[7] See also Chomsky 1998.

This "array of lexical choices" is computed upon by C_{HL}, the operations of which "recursively construct *syntactic objects* from items in N and syntactic objects already formed" (ibid: 226). The operations of C_{HL} are divided into three components. The one which first applies to N is the OVERT COMPONENT. This overt component includes the operation SELECT, which picks an LI from N -- thereby reducing its index by 1 -- and introduces it into the computation as a syntactic object. There is also an operation, MERGE, which "takes a pair of syntactic objects (SO_i, SO_j) and replaces them by a new combined syntactic object SO_{ij}" (ibid). For the computation as a whole to be a DERIVATION -- i.e. for the computation to actually yield an SD -- Select must reduce N to zero, and Merge must apply often enough to give a single SO at the interface levels.

The overt component computation "splits into two parts" (ibid: 229), one carried out by the COVERT COMPONENT and the other by the PHONOLOGICAL COMPONENT. This "split" occurs at the application of the operation SPELL-OUT, which "strips away from Σ [the structure formed by the overt component] those elements relevant only to π" (ibid). These phonological elements are mapped by the phonological component to π at PF. The phonological component also has a morphological subcomponent which "constructs word-like units" (ibid).[8] What remains of Σ is mapped by the covert component to λ at LF.

At the interfaces, PF and LF, the computation may CONVERGE or CRASH (ibid: 220). It *converges* if the representations it yields, π and λ, satisfy the interface convergence condition FULL INTERPRETATION (FI), and *crashes* if they do not. FI states that π must consist entirely of LEGITIMATE PF OBJECTS, and λ of LEGITIMATE LF OBJECTS (ibid: 194, 219). These legitimate PF and LF objects give INSTRUCTIONS -- consisting of information "about the properties of each linguistic

---

[8] "Word-like units" is a rather enigmatic phrase, and one that prompts the question: how exactly are (a) word-like units, (b) words, (c) SDs, (d) lexical items, and (e) syntactic objects different from, and related to, one another?

expression, including its sound and its meaning" (ibid: 21) -- to the A(rticulatory)-P(erceptual) and C(onceptual)-I(ntentional) performance systems respectively. The A-P and C-I systems in their turn assign INTERPRETATIONS to the FI-satisfying π and λ representations.

To summarise: an I-language generates an SD when its C_{HL}'s computations upon an N, an array of lexical choices, constitute a convergent derivation -- i.e. a derivation that is forms, at PF and LF, a (π, λ) pair which gives appropriate instructions to, and assigned appropriate interpretations by, the appropriate performance systems.

2.1.2 Chomsky's internalist semantics

λ, LF representation of the SD pair, is a symbolic object which consists of "semantic features". According to Chomsky:

> We may take the semantic features *S* of an expression *E* to be its *meaning*... *E means S* in something like the sense of the corresponding English word,... *S*... providing the relevant information for the performance systems. (1995a: 19-20)

So it is λ that appears to be the locus of the I-language-generated LE's semantics, and therefore central to Chomsky's account of linguistic semantics. The important point is that this account is part of an internalist theory -- one that holds that language is an innate, internal, mentally-constituted component which "produces" LEs only insofar as it specifies or determines their properties. The LEs of Chomsky's linguistic theory exist and have certain properties only in virtue of being specified or described by a component -- i.e. the I-language -- in the human mind/ brain, and not because they are or can be uttered by

humans.[9] Similarly, these LEs have certain semantic properties M because the I-language states that they do -- not because people utter or interpret them to mean M.[10]

Hence Chomsky's condition of INCLUSIVENESS:

> A "perfect language" should meet the condition of inclusiveness: any structure formed by the computation (in particular, π and λ) is constituted of elements already present in the lexical items selected for N; no new objects are added in the course of computation apart from rearrangements of lexical properties... (ibid: 228)

In other words, the π and λ generated by the "perfect" language ought to "consist of nothing beyond properties of items of the lexicon (lexical features)..." (ibid: 225).

The lexical properties that form λ constitute

> ... a rich texture of purely internalist semantics, with interesting general properties, and evidence for formal semantic relations (including analytic connections). (1992b: 218)

-- a semantics which arises from, or is based on

> ... a rich conceptual structure determined by the initial state of the language faculty (perhaps drawing from the resources of other genetically determined faculties of the mind), waiting to be awakened by experience. (1992a: 115-6)[11]

---

[9] This is not to say that I-language is peculiar to humans.

[10] All this, of course, is on the assumption that mentally-constituted, truly *I*-linguistic expressions can actually be "*uttered*".

[11] The possibility of the language faculty "drawing from the resources of other genetically determined faculties of the mind" (1992a:115) raises questions concerning the nature of Chomsky's supposedly internalist *linguistic* semantics. He acknowledges that "Questions remain... as to just what kind of information is within the lexicon, as distinct from belief systems" (1992b: 217); and also asks, "In what respects... do these [linguistic semantic] properties belong to the language faculty as distinct from other faculties of mind to which it is linked? How do lexical resources relate to belief systems, for example...?" (1995a: 23). It seems to me that these questions are important -- not because they enrich, but because they undermine, Chomsky's minimalist program. The notion that language "draws from" other cognitive systems for its semantics, and that lexical items are "*based on* conceptual structures" (1992a: 113, my italics), implies that language is

This is why, Chomsky argues, "it is a very difficult matter to describe the meaning of a word" (ibid: 113). Notions of location, actor, recipient, event, intention, etc, together with their "specific properties and interrelations", "enter widely into lexical structure, often in quite abstract ways" (ibid.). 'Chase', for example, is an apparently simple word which actually has a highly intricate meaning:

> To chase Jones is not only to follow him but to follow him with the intention of staying on his path, perhaps to catch him. (ibid.)

But the important point is that this semantics, though "rich" and "intricate", is nevertheless "purely internalist", "determined by the initial state of the language faculty", and has no direct relation to the entities and situations in the world (or any actual or possible world) which it supposedly "describes" or "represents". λs (and πs) are mental representations, symbolic objects --

> ... postulated mental entities, to be understood in the manner of a mental image of a rotating cube, whether the consequence of tachistoscopic presentations or of a real rotating cube or of stimulations of the retina in some other way; or imagined, for that matter. (1995a: 53)

-- and Chomsky reminds us that "all of this is pure syntax, completely internalist..." (ibid: 19). LEs with λs do not represent the world, or pick out things in the world. They are not related to the world in any Fregean or Peircean sense. Such relations, Chomsky argues,

---

primarily to do with the *vehicle* half of the vehicle-meaning relation, linking π/ PF properties to concepts that are essentially pre- and (therefore) non-linguistic in nature. In other words, the implication is that λ and π are not on an equal linguistic footing, the semantic features which λ consists of being "linguistic" only to the extent of having been assigned a vehicle by I-language. This idea is reinforced by Chomsky's suggestion "... that there is something like an array of innate concepts and that these are to a large degree merely "*labeled*" in language acquisition..." (ibid: 116, my italics). See chapter 5.

are invented technical notions: terms like *extension*, *reference* and *denotation* are "technical innovations, which mean exactly what their inventors tell us they mean" (ibid: 42), and "[have] no counterpart in ordinary language" (ibid: 24).

For Peirce (1933), reference is a triadic relation between person X, real object in the world Y and sign S, whereby X refers to Y by S. Frege suggests that

> A logically perfect language (*Begriffsschrift*) should satisfy the conditions, that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact designate an object, and that no new sign shall be introduced as a proper name without being secured a reference. (1952: 70)[12]

Both Peirce and Frege are mainly concerned with "signs" having *truth value*, i.e. being true or false of objects in the world (in Peirce's case, real objects in the real world). Chomsky, on the other hand, seems to be denying that LEs have any such relation with the world in virtue of their semantic properties.[13] All they do, he claims, is to *provide perspectives for viewing things*:

> Perhaps the weakest plausible assumption about the LF interface is that the semantic properties of [an LE] focus attention on selected aspects of the world as it is taken to be by other cognitive systems, and provide intricate and highly specialized perspectives from which to view them... (1995a: 20)

> ... a lexical item provides us with a certain range of perspectives for viewing what we take to be things in the world, or what we conceive in other ways; these items are like filters or lenses, providing ways of looking at things and thinking about the products of our minds. (1992b: 221)

---

[12] By "proper name" Frege means any designation which has as its reference a definite object (1952: 57).

[13] This view of LEs and their semantics is not entirely at odds with Peirce's definition of reference as a *triadic* relation. It could be argued that Peirce's sign S, in itself, does not refer to anything either -- it requires *person X* to *use it* to refer to Y (real object in the world).

That is, Chomsky's LEs do not refer to things, but merely "provide ways of looking at things". Furthermore, the *things* which they provide perspectives for viewing are most emphatically NOT *things in the world*, but "*selected aspects of the world as it is taken to be by other cognitive systems*", or "*what we take to be things in the world*". In other words, what is viewed through the LE-supplied perspectives is not the world itself, but our conception of the world -- or, more specifically, the particular perceptual/ cognitive effects, in particular circumstances, which the world produces in us.

So there are two points which we have to bear in mind. The first is:

> (1) it would appear that Chomsky's phrase "the products of our minds" -- these "products" including, of course, the world *as we perceive/ cognise it* -- covers *everything* which purely internalist LEs are in some way semantically related to.

The second (and probably more important) point is that

> (2) this relation is as necessary to the existence and nature of the internalist semantics as, say, light waves are to the existence and nature of "filters or lenses" -- that is to say, *not necessary at all*.

In general, Chomsky rejects the notion of a "common public language" -- like Frege's *Begriffsschrift* -- consisting of "grammatically well constructed" signs that have a fixed relation ("reference") to objects in the world taken as things divorced from the intrinsic properties, and from particular speakers' use, of the mode of designation (i.e. the signs themselves) (1992b: 222, 1995a: 26). He also disagrees with Dummett's suggestion that a speaker-hearer's linguistic ability consists in a "partial, and partly erroneous, grasp" (Dummett, 1986: 468) of some common public language whose words have meanings "independently of any particular speakers" (ibid: 473). Chomsky's position can be

explained in terms of both (1) and (2). He is against the idea of a reference-based common public language, not only because he believes that LEs have a semantics regardless of the things in the world that they are said to "represent" ((2)); but also because it is arguably only through the perceptual/ cognitive systems of "particular speakers" that aspects of the world are available for the signs of a common public language to be related to ((1)).

I feel it is necessary to underline (1) and (2), for they have implications which Chomsky does not unequivocally set out, and whose place within his internalist/ minimalist framework he does not clearly indicate. To reiterate: (2) has to do with Chomsky's own remarks about a linguistic semantics providing nothing more than filters or lenses to view things through. It follows from this that the things, and the viewing of them, have nothing to do with the existence and nature of the semantics itself. Point (1) is that the things which we view through the semantically-supplied perspectives are NOT *things in the world*, but *what we take to be things in the world*. In other words, Chomsky's linguistic semantics is the filters/ lenses through which we can view *our view of the world* -- i.e. not the world itself, but our (presumably) pre-linguistic view of the world. So, taking (1) and (2) together, it could be said that language for Chomsky is at two removes from objects in, and states of, the real or some possible world, since it is

> (1') *the world as we perceive/ cognise it*, not the world itself, that is
>
> (2') *viewed through the perspectives provided by the linguistic semantics*, rather than directly referred to by LEs.

And if language within Chomsky's internalist framework is at *two* removes from the world, we should also remember what it is at *one* remove from -- (1'), or "the products of our minds" (Chomsky 1992b: 221). In other words: while LEs are not directly related to things in the world, neither do they appear to be directly related to "the products of our

minds" which Chomsky's linguistic semantics provides perspectives for viewing.[14] Or rather: LEs represent neither the world nor our thoughts, impressions and experiences of the world.

This brings me to an important point which I feel is insufficiently stressed by Chomsky: that, not only is it true that LEs do not refer to objects in the world, they do not even "express shared thoughts" (1995a: 26) along the lines of Frege's *Sinn*. Perhaps they cannot be said to *express* anything at all, for it follows from Chomsky's account of internalist semantics that his LEs do not have the sort of truth-conditional "meaning" that belongs to conscious thought or communicative utterance. What they do have are λs -- the study of which is, as Chomsky puts it, pure syntax, since λs are nothing more than mental representations or symbolic objects.

Chomsky does use the phrase "the meaning of a word" (1992a: 113). And, in the quote (1995a: 19-20) which I use at the beginning of this section, he suggests that the semantic features *S* of expression *E* consititute *E*'s *meaning*. But he goes on to say, rather equivocally, that "means" in *E means S* is to be taken "in *something like* the sense of the corresponding English word" (my italics). As applied to *E* and *S*, what Chomsky means by "mean/ meaning" cannot be anything like "the sense of the corresponding English word" if it is to be consistent with (1), (1'), (2) and (2'). Such a λ-based, perfectly internalist linguistic semantics as he outlines is simply and inherently NOT expressive of propositions which (i) individuals may mentally entertain or physically communicate; and (ii) are, in virtue of having (or being able to have) truth value, mentally entertained or physically communicated for the purpose of representing states of the world. Henceforth I shall use "meaning" to refer to propositions expressed and entertained (and sought for) by human

---

[14] These "products of our minds" are, presumably, the products of cognitive systems *other than the I-language*: "... the semantic properties of [an LE] focus attention on selected aspects of the world as it is taken to be by *other cognitive systems*..." (1995a: 20, my italics).

individuals, and something more specific like "(linguistic) semantic content" for the semantic elements of Chomsky's I-linguistic expressions.

It is possible that the features of Chomsky's internalist semantics mentioned above -- as specified by, or deducible from, his own remarks -- are part of what makes his notion of language INCLUSIVE as well as INTERNALIST. In other words, these may be the properties that define inclusiveness as well as internalism with regard to the λ/ LF side of I-language. Chomsky's LEs are *internalist* in that they are *not* a set of utterances or behaviours which include the act of verbally referring to objects in the world, and *semantically internalist* in that there is nothing in their semantics that allows them to refer (or even to be used to refer) to objects in the world. And, leaving aside objects in the world and taking just the mind into consideration, it could be that the *semantic aspect of inclusiveness in general* -- i.e. LEs being generated from nothing but I-linguistic lexical items -- consists in the basic irrelevance to, and exclusion from, linguistic semantics of other "products of our minds".

To summarise: in virtue of its internalist and inclusive nature, Chomsky's linguistic semantics is not a relation between LEs and objects in the world ("reference"). It is not even a relation between LEs and the thoughts of an individual with an LE-generating I-language, in that it is very far from being a set of "meanings" which are intrinsically and directly expressive of the individual's percepts, knowledge and intentions. Minimalist semantic properties merely provide perspectives -- albeit very intricate ones -- through which things (i.e. "the products of our minds") may be viewed.

2.1.3 The "USE" of (I-)language

Although

(3) the semantic content of Chomskian LEs

is by its nature nothing like, and completely independent of,

(4) the sort of propositional meaning which humans are concerned with;

the fact is that (4), according to Chomsky, may be expressed through the USE of (3). Chomsky claims that the process of using LEs involves the performance systems which interface with the I-language, as well as other conceptual resources (e.g. belief systems) to do with "human interests, intentions, goals and actions" (1995a: 30). It is these cognitive systems, and not the linguistic semantics, which are largely responsible for -- which provide the crucial, expressive, truth-evaluable part of -- the "meanings" that we can express through the USE of language. Even the very elementary notion of *nameable thing*, Chomsky argues, is specified not by language, but by *us*. He points out that

> What we take as objects, how we refer to them and describe them, and the array of properties with which we invest them, depend on their place in a matrix of human actions, interests, and intent... (1992b: 207)
>
> Even the status of (nameable) thing, perhaps the most elementary concept we have, depends crucially on such intricate matters as acts of human will... (1995a: 22)

For example, a person who comes across some sticks on a road might not take them to be a thing at all, unless he discovers that

> ... they were specifically constructed as some kind of object, whether by people or, perhaps, beavers. (1995a: 30)

So *nameable thing* is not so much *a* concept as *our* concept; it is inseparable from the "human interests, intentions, goals and actions" which form the basis of, or perhaps constitute, our conceiving of it.

Chomsky is careful to distinguish such concepts from the semantic content of internalist LEs. For him, the notion of nameable thing is a part of "common sense understanding" or "how people interpret object constancy, the nature and causes of motion, thought and action, and so on" -- the study of which he calls "folk science" (1995a: 28). He also uses the term "ethnoscience" to refer to the more general study, encompassing "folk science", of "what people take to be constituents of the world, however they may talk about it" (ibid: 30). The important point is:

> *The study of semantic resources of the language faculty is not ethnoscience...* (ibid: 31, my italics)

Chomsky claims that it is the cognitive systems and processes involved in "common sense understanding" -- NOT an internalist semantics -- which determine concepts like that of nameable thing; and that these concepts are, consequently, the object of ethnoscientific rather than linguistic inquiry.

It is these concepts which constitute (4) from the beginning of this section, i.e. the sort of propositional meaning which humans are concerned with -- and which, Chomsky argues, may be expressed through the USE of LEs. Therefore, if Chomsky's "study of semantic resources of the language faculty" is not ethnoscience, then neither is it the study of *how LEs are used*. Nor is it the study of *how utterances are interpreted*; or rather, how utterances are produced to express, and interpreted as expressing, (4) -- utterance production/ interpretation being (to my knowledge) the one clear example of language use that Chomsky provides. Chomsky (1992b, 1995a) agrees with Davidson's view that knowing how to interpret utterances is indistinguishable from "knowing our way around in

the world generally" (Davidson 1986: 446), and points out that utterance interpretation therefore involves not just the language faculty but also "all other capacities of the mind, whatever they may be" (1992a: 120).

Note that Chomsky seems to take it for granted that *to utter LEs* is *to use them for communication*, for he asserts that:

> The study of communication in the actual world of experience is the study of the interpreter [of utterances]... (1992a: 120)[15]

Moreover, *communication* encompasses the rather more diffuse list of uses for language -- all of them to do with utterance production/ interpretation -- which he mentions elsewhere:

> ... articulation, perception, talking about the world, asking questions, telling jokes, and so on. (1995a: 12)

Thus Chomsky also claims that "the study of communication in the actual world of experience" -- including, of course, the study of how LEs are used in the production and interpretation of utterances -- is actually "the study of everything" (1992a: 120). I take "everything" in this context to mean "knowing our way around in the world generally", "all other capacities of the mind", and "human interests, intentions, goals and actions".

For instance, the notion of nameable thing may be communicated or retrieved when certain LEs (e.g. 'nameable thing', 'cat', 'The cat sat on the mat', 'Go and catch a falling star') are used. For Chomsky, what is communicated/ retrieved is not contained within those LEs (at least not completely), and has a great deal more to do with how and why the speaker used the LEs, and (perhaps more importantly) with the interpreter's assumptions about how and why the speaker used the LEs -- all of which, of course, is based on the

---

[15] In Chomsky 1971 (: 19), however, he remarks that language use "need not involve communication, or even the attempt to communicate". See also section 4.4.

"human interests, intentions, goals and actions" of the speaker and hearer, and what they take to be the constituents of the world.

It has to be pointed out that, although he discusses language use in terms of *utterances* produced by *speakers* (e.g. in Chomsky 1995a and 1995b), Chomsky's (and Davidson's) theories and comments should apply equally to, say, orthographic inscriptions produced by a writer. For convenience, I shall use the term "UTTERANCE" to cover all E-physical phenomena that humans produce in virtue of "using language" -- speech, writing, the gestures of sign language, etc. The issues of how these different types of E-physical phenomena are related to one another, and whether *all* of them can really be accounted for by various important theories of language and language use, are important ones. However, my present concern is only with the physical nature of utterances, and its implications for Chomsky's notion of language use.

For Chomsky, the use of LEs in humans is in part effected by the performance systems, which include the articulatory-perceptual (A-P) systems. Also, the I-linguistic phonological component which interfaces with the A-P systems has corresponding "special properties" that "relate to the need to produce instructions for the sensorimotor systems, for production and perception" (1995b: 229). Assuming that it is impossible for a Chomskian derivation to bypass PF (much less the phonological component within the I-language),[16] and bearing in mind that PF is the interface with the *articulatory-perceptual* systems, it would appear to follow from Chomsky's own conception of the language faculty that the process of language use for human beings *must* in some direct or indirect way involve the articulation and perception, the production and interpretation, of articulable/ perceivable -- i.e. *E-physical* -- utterance phenomena.

---

[16] That a derivation cannot bypass PF does not actually need to be *assumed*, since it is *presupposed* by the very concept of "derivation" -- according to Chomsky, a derivation is a derivation (rather than just a computation) iff it converges at PF (and LF).

It may seem superfluous to stress that "language use" as discussed throughout this chapter is *language use in humans*. However, Chomsky claims that

> Some other organism might, in principle, have the same I-language (brain state) as Peter, but embedded in performance systems that use it for locomotion. (1992b: 213)

-- so I feel it is necessary to point out that "language use" in the strictest Chomskian sense, while not necessarily a process which distinguishes humans from other species, has to do with utterance production and interpretation if and only if the I-language is embedded in the appropriate sort of performance systems, i.e. the sort which humans possess. Put another way: within Chomsky's framework, humans may not be the only organisms to have, and use, I-language; but it is possible that *only* human I-language is embedded in performance systems designed for the production and interpretation of E-physical utterance phenomena.

The important point, however, is that these performance systems are *A-P* as well as C(onceptual)-I(ntentional). While abstract concepts and intentions are clearly required for what Chomsky regards as language use in humans, it seems that the E-physical, and the articulation and perception of the E-physical, are also an integral part of the process. All this implies that human I-language, according to Chomsky, is used primarily for the production/ interpretation of E-physical utterance phenomena, and for communication in particular. I say "primarily" rather than "only" for utterance and communication, for Chomsky's theory does appear to allow for the possibility of "thinking in LEs". Although language tends to be very strongly associated with communication (or at least with utterances), there is no doubt that we do "think in LEs" -- that, among the different kinds of thought or the various ways in which we think, there is at least one which language is a crucial part of. But because the human I-language is specifically and inextricably embedded in performance systems that have to do with utterance production/

interpretation, and because these performance systems effect *all* uses of the I-language, it would appear that language use in humans (as conceived by Chomsky) *must* somehow involve utterance phenomena, whether a person is speaking to someone or "thinking in LEs". It follows from this that, within Chomsky's framework, the LEs one thinks in can only be in the form of *internalised utterances*, i.e. mental representations of physical forms (or potentially physical forms) defined, unavoidably, by PF and the A-P systems.

So an example of the only sort of "thinking in LEs" permitted by Chomsky's theory of language is when a person *thinks up LEs to utter* -- i.e. what goes on in one's brain when one is framing expressions for writing down or saying aloud; like what I am thinking now as I write these words. This kind of mental use of LEs is nevertheless closely related to communication, for it is aimed at the formation of expressions that are *for uttering*.[17]

---

[17] Of course this "thinking in LEs" need not necessarily be the thinking up of LEs to utter to someone -- the internalised utterances may very well have been directed at oneself. As may be the case even for entirely E-physical "uninternalised" utterances, although I have been assuming that to utter LEs is to use them for communication. Black (1979: 188-9) points out that the "soliloquizing thinker" presents a problem for Searle's (1969, 1975) speech-act approach:

> Any clear cases of speech-acts that come readily to mind involve communication with an audience: it makes little sense to think of promising *oneself* something, or warning, advising, pronouncing judgement, and so on, to oneself. (Black 1979: 188)

In my view, there is a similar problem for those theories of language use according to which certain meanings are attributed to utterances of LEs on the basis of assumptions that two or more conversational participants are exhibiting cooperative behaviour (Grice 1975), or that utterances are intended by their producers to convey relevant information to the hearers (Sperber and Wilson 1986a/ 95). Such assumptions are clearly unnecessary in the case of a person speaking to himself, since that person is privy to his own intentions. So the phenomenon of *speaking to oneself* seems devoid of purpose, and consequently inexplicable, if considered within theories of *speaking to someone else*. But perhaps it is not "speaking to yourself" which is *pointless* in terms of "speaking to someone else", but "speaking to someone else" which is *impossible* compared to "speaking to yourself". After all, assumptions of cooperation and relevance are merely assumptions; and there is no guarantee that an interpreter ever interprets correctly or that a speaker ever communicates perfectly. Strictly speaking, the speaker/ interpreter only has his own experiences to go on, rather than any true knowledge of other people's states of mind. In attempting to communicate something clearly and unequivocably to Mary, for example, Peter's expectations concerning Mary's behaviour and interpretive abilities have much less to do with what *she* actually knows and is capable of, than with what *he thinks* she knows and is capable of. This, in turn, is based on what he thinks the process of utterance interpretation requires; indeed, what he himself would apply to the process

Note once again that "thinking in LEs" or thinking which involves language (whether or not *language* here is Chomsky's I-language) is merely one of several different kinds of thought. Note also Chomsky's position in respect of the debate over whether language is necessary for thought or vice versa: while he says nothing about "thinking in LEs", it is clear that he does subscribe to the notion that there is some variety of pre- and non-linguistic thought which, furthermore, is a prerequisite for both I-language and language use. From the "a priori framework of human thought, within which language is acquired" (1992a: 114) and the "conceptual structures" (ibid: 113) or "innate concepts" (ibid: 116) upon which I-linguistic lexical items are semantically based, through the numerations assembled from LIs according to the "choices and intentions of speakers" (1995b: 237), to the "human interests, intentions, goals and actions" which determine the production/ interpretation of utterances -- there is a sense that certain mental processes and entities drive the use (via performance systems) of I-language, as well as the mechanisms of the I-language itself. However, these (i) non-linguistic mental processes/ entities and their role in the workings of the language faculty must not be confused with (ii) the sort of mental activity which *does* involve I-language, and the part which I-language in its turn plays in this process of "thinking in LEs". And of course it is (ii) which I have been discussing -- my argument being that, if LEs are pairs of PF and LF representations generated by a Chomskian I-language, then "thinking in LEs" would be equivalent to the *use* of LEs for thinking.

Taking into consideration Chomsky's view that the performance systems in which the human I-language is embedded consist of A-P as well as C-I systems, I also argue that this *use of LEs in thought* must be secondary to, and more circuitous than, *the use of LEs in the production and interpretation of utterances.* This is because all uses of LEs are

---

were he the interpreter. On Mary's side, it could be said that her beliefs about speaker behaviour, which determine the assumptions she makes about Peter's intentions, are founded on her own experiences, intentions and behaviour as a speaker.

(according to Chomsky) facilitated by the same performance systems, the A-P elements of which would automatically convert LEs into some "utterable" form even if the intention was to use the LEs for thought rather than in utterance.

As I am mainly interested in the semantic aspects of Chomskian linguistics, I began this section on language use with a discussion of the discrepancy between (a) the (purely linguistic) semantic content of LEs and (b) the propositional meaning which LEs can be used to express. From this semantic point of view, it could be said that the most important elements of Chomsky's theory of language and language use are (i) λ; (ii) the half of the language faculty leading to, and partly constituted by, LF and the C-I performance systems; and (iii) how a λ emerging from the covert component of the I-language is endowed, by the C-I systems (and other conceptual resources like belief systems) at the LF interface, with the propositional meaning that it is used to express.

However, it also follows from Chomsky's theory that the propositional meaning which arises from the use of the *human* I-language is not simply EXPRESSED. A fuller characterisation of this type of propositional meaning, along Chomskian lines, is that it is COMMUNICATED -- or *can* be communicated -- and communicated, specifically, by means of particular E-physical phenomena (i.e. UTTERANCES) to other non-telepathic humans. In that case, what the C-I systems do to λ is merely part of a process whose uniqueness lies in the nature of the *entire* set of the performance systems which effects that process -- i.e. not just the C-I systems, but also the *A-P* systems, whose operations on π determine the E-physical features of language use. In short, it seems that language use in humans as defined by Chomsky occurs only when LEs interact with BOTH types of performance systems. (This is anyway indicated by the fact that the Chomskian LE is an SD consisting of a λ that gives instructions only to the C-I systems, and a π that gives

instructions only to the A-P systems -- so that, for the whole LE to be used, instructions have to be taken from both λ and π, and therefore by *both C-I and A-P systems.*)

I conclude this section with some general remarks on Chomsky's concept of language use; and especially on how it brings into relief what are probably the most important and fundamental features of the object which is used, I-language.

There are two main points. The first has to do with the apparent INACCESSIBILITY of the I-language to the human consciousness in the absence of the performance systems. The phrase "language use" may give the impression that the I-language is something which the human individual H is conscious of possessing; whose capacity for being used for communicating propositional meaning H is intelligently aware of; and which H *consequently* uses for that purpose. But the converse appears to be true: it is not that H uses I-language and LEs because he is conscious of their existence, but that he is conscious of them only insofar as he does or can use them. According to Chomsky, there are certain things which a person -- say, Jones -- knows in virtue of having an I-language. However,

> ... Jones has this knowledge whether or not he is aware of these facts about himself [i.e. the fact that he has an LE-generating I-language, and "facts" about the I-language and the LEs it generates]; it may take some effort to elicit such awareness, and it might even be beyond Jones's capacities. This is a question that falls within the broader context of performance systems. (1995b: 17)

Here, Chomsky seems to be saying that it is only through certain performance systems that Jones is aware of having an I-language -- the degree and quality of his awareness depending on the character of his performance systems. Another way of putting this is that any "awareness" of his I-language which Jones may have would be *constituted* by the interaction between his performance systems and I-language -- i.e. by the process of language use as carried out by the performance systems.

Thus, given that

(5) our performance systems, and therefore our use of language, provides our consciousness with the only route to our I-language; and that

(6) these performance systems have to do with utterance production/ interpretation,

Chomsky's claim appears to be that it is only in virtue of producing and interpreting utterances, or at least of being able to produce and interpret utterances, that we ever gain conscious mental access to I-language and the LEs it generates.

But this claim -- that it is the performance systems which establish the I-language in the very prominent place it undoubtedly occupies in the human consciousness -- must not be allowed to obscure the fact that the "I" in "I-language" also stands for "INTENSIONAL". In other words, it has to be balanced with what is perhaps the most fundamental of Chomsky's claims about the nature of language: that it is *intrinsically* I-linguistic, and *not* defined by what it does or what happens to it -- or, in the case of humans, what it is used for. Although under certain circumstances it can be used for utterance production/ interpretation, Chomsky's I-language is not essentially a facility or instrument for communication. As he puts it:

> In general, it is not the case that language is readily usable or "designed for use." The subparts that are used are usable, trivially; biological considerations lead us to expect no more than that. (1995b: 18)

So, its very close and important relation to the performance systems notwithstanding, the (human) I-language is not "for" the production/ interpretation of utterances, or the

communicating of propositional meaning. In fact no I-language, human or otherwise, ought to be *for* anything.

These, then, are my two points:

(7) On the one hand, the performance systems and the I-language are (at least in humans) so closely related that, without the former, the latter could be said to be of no significance whatsoever to the conscious human individual. Or, to put it another way, the human individual must, as it were, take the performance systems with the I-language if he is to "have" the I-language at all.

(8) On the other hand, Chomsky's intensionalist notion of language has it that the I-language is defined by inherent properties that are obviously not determined or affected by anything external to it -- not even by the performance systems which it is embedded in, and interfaces with.

In semantic terms, (7) and (8) may be expressed thus:

(7') We tend to think that LEs have fully propositional, mentally entertainable meanings; and that in these meanings lies the significance, the primary epistemic value, of language for us. Or rather, what we refer to, broadly, as "the meanings of words/ sentences" is what we generally believe to be the whole "point" of language (even though there is no real reason for language to have anything like a "point" or a "purpose"). But according to Chomsky, such meaning only arises through the use of -- via the interaction of the C-I systems with -- the I-language. Thus Chomsky's claim is that, without the C-I systems, what we regard as central to language would be lost to us.

(8') Nevertheless, we must bear in mind that propositional meaning is not expressed by the I-language, but through the USE of the I-language -- the process and products of which are not supposed to have any bearing on the intrinsic nature of the I-language itself. Such

propositional meaning, therefore, *does not constitute a linguistic semantics*. For example, it is obviously a matter of language use, and not one of (I-)linguistic semantics, that the I-language is used to describe or represent the world. Or, as Chomsky points out, his internalist/ intensionalist framework has

> ... no provision for what Scott Soames calls "the central *semantic* fact about language,... that it is used to represent the world", because it is not assumed that language is used to represent the world, in the intended sense... (1995a: 27; Soames 1989; my italics)

And if the USE of language to represent the world, or to express propositional meaning in general, is *not* a LINGUISTIC SEMANTIC fact, then:

> It is possible that natural language has only syntax and pragmatics; it has a "semantics" only in the sense of "the study of how this instrument, whose formal structure and potentialities of expression are the subject of syntactic investigation, is actually put to use in a speech community"... (ibid: 26)

What Chomsky seems to be saying is this: if the term "semantics" has to do with the sort of propositional meaning which the conscious human mind is interested in, then I-language does not have a semantics, because that sort of meaning is not expressed by the I-language itself but only through the use (by an intending agent) of the I-language.[18]

---

[18] However, as Chomsky himself uses the terms "semantic features/ properties/ relations" and "internalist semantics" in his discussions and descriptions of his I-language, I shall continue referring to λs and LF objects as the "semantics" of LEs.

2.2 How exactly does the I-linguistic semantics bear upon the "literal use" of an I-language?

In chapter 1 I argued that an account of what is metaphorically expressed by an utterance of, say, the LE 'Bill is a bulldozer' must include an account (or at least an assumption) of what that utterance literally expresses. Furthermore, I claimed that literal meaning[19] is a USE notion, albeit one which presupposes that the *vehicle-meaning pairs* used to express literal meaning are antecedent to, and independent of, *how they are used*. Thus literal meaning is conveyed by the literal use of a vehicle of meaning. The literal use of such a vehicle is to use it to mean what it autonomously means, rather than to mean what one wants it to mean. In chapter 1 I also argued that literalness tends to be associated with the notion of language, for the reason that the latter is often assumed to be a fixed set of rules determining autonomous vehicle-meaning relations, and therefore a suitable locus for a stable, accessible literal meaning upon which metaphorical meaning may be constructed.

Note that my distinction between vehicle-meaning relation and the use of that relation does not necessarily entail a distinction between *the meaning expressed by the vehicle*, on the one hand, and *the meaning expressed by the literal USE of the vehicle*, on the other. However, that Chomsky's I-linguistic semantics *cannot in fact be regarded as literal meaning* is evident from two different but related aspects of his theory of language: (i) the nature of the "meanings" carried by the vehicles which constitute Chomsky's LEs, and (ii) Chomsky's conception of language use. With respect to (i): within Chomsky's internalist framework, the autonomy of the linguistic vehicle-"meaning" pairs (i.e. the PF and LF representations) appears to follow from, or to take the form of, an ontological distinction between the "meanings" of LEs and the sort of propositional meaning which is expressed through the use (literal or otherwise) of LEs. I have been putting quotation marks around

---

[19] As I have stated in section 2.1.2, I am using "meaning" (e.g. "literal meaning", "the meaning of expressions") in the sense of *propositional meaning which humans are interested in*.

"meaning" connected with Chomsky's LEs, because these LEs *do not have propositional meaning* -- the supposedly internalist and inclusive linguistic semantics merely provide perspectives for viewing the world. Thus the *non-propositional* linguistic semantics of Chomsky's LEs obviously cannot be the *propositional* literal meaning of those LEs.

Apart from the fact that Chomsky's linguistic semantics does not constitute the sort of propositional meaning which may be regarded as literal, there is also (ii): Chomsky's notion of the use of his LEs. Literalness is not just an expression meaning what it means, but rather an expression being interpreted, or intended to be interpreted, as meaning what it means. In other words, literalness involves conscious thought, intention -- which, as far as Chomsky's I-linguistic expressions are concerned, is only achieved through the *performance systems* that facilitate their use.

Thus, in his linguistically-orientated version of the use of the vehicle-meaning relation, Chomsky sharpens and makes absolute the distinction between vehicle-meaning pair and the use of the vehicle: by attributing to the linguistic "vehicles" a non-propositional, internalist and inclusive linguistic semantics, and by relegating the use of the linguistic vehicle-"meaning" pairs to external, extra-linguistic performance systems. While this has the effect of emphasising the distinction, it also raises the following questions:

(9) How exactly does the use of the linguistic semantics involve or implement the linguistic semantics?

(10) How is the propositional meaning which is expressed through the use of the linguistic semantics actually related to the non-propositional linguistic semantics?

Chomsky does assume that how an I-language is used is at least to a certain extent constrained by the I-language itself. What is more, he seems to be of the opinion that each I-language corresponds to a particular use of that I-language to convey a particular set of

meanings -- a use and a set of meanings which may be referred to as literal. From this quote --

> Jones may speak in a way that is not in accord with his I-language, or may offer judgements inconsistent with his I-language... (Chomsky 1992a: 122)

-- it can be deduced that there is supposed to be one particular use of Jones' I-language, expressive of a particular set of meanings, which is "in accord" with the I-language, and which (I argue) is what Chomsky would regard as *literal*. But Chomsky does not provide any answer to the question of the relation between I-language(/I-linguistic semantics) and language use(/propositional meaning expressed through language use). In fact he appears to evade the question by contradicting his own internalist conception of language. For a start, how does Jones use his LEs in a way that is "in accord with his I-language", when Chomsky's view of language is such that the language-user has no real idea of the nature of his innate, internalist and inclusive I-language? As I pointed out in the previous section, Chomsky suggests (1995b: 17) that an awareness of having an I-language, and of the LEs generated by the I-language, may be beyond the capacities of Jones.

More importantly, Chomsky also argues that the question of whether or not Jones is aware of having an I-language "falls within the broader context of PERFORMANCE SYSTEMS" (ibid, my caps). According to Chomsky, a person is aware of having an I-language only in virtue of being conscious of uttering LEs, or in virtue of his capacity for uttering LEs. He does not use I-language because he knows he has one -- rather, he knows he has an I-language in virtue of using it. Thus a human's awareness of his I-language is constituted by his use of it. This means that he may be conscious that *he* is using his I-language, but there is no reason to suppose that he knows what it is that he is using -- or, more specifically, what it is that his performance systems are accessing.

Because our relation to I-language is always mediated by the performance systems, we can only be conscious of LEs, not "as they are", but only as they are used by us.

Yet Chomsky claims that Jones

> ... may *choose to violate* the rules [of his I-language]... (1995a: 36, my italics)

This clearly implies that "the rules of an I-language" are things which Jones may consciously, volitionally and directly adhere to or deviate from.

The same sort of contradiction occurs in Chomsky's use of the phrase "knowledge of language" (as I have already mentioned in section 2.1, fn1). In Chomsky 1965 (: 4) he refers to the "the speaker-hearer's knowledge of his language" as "linguistic competence", and simultaneously argues that language is constituted by this "competence". As recently as 1995 (b: 14), Chomsky equates "mature linguistic competence" with the steady state I-language. He also describes "competence" (ibid) as the speaker-hearer's "knowledge and understanding", and contrasts it with "performance", i.e. "what [the speaker-hearer] does with that knowledge and understanding". Thus "knowledge of language" for Chomsky cannot be "knowledge" in the usual relational sense of *conscious, and consciously acquired, awareness OF language*. As he himself puts it, "knowledge of language" is "knowledge without grounds" (1986: 12).[20] This "knowledge of language" simply *is* language -- is, in fact, *constitutive of internalist and inclusive I-language*.

But Chomsky also uses "knowledge" in "knowledge of language" in the relational sense, conflating I-linguistic semantics which is supposed to be (partly) constituted by "knowledge of language" in the non-relational sense, with the conscious and relational "knowledge" which the speaker-hearer has of what particular LEs can be used to expressed. For example, he claims that

---

[20] See also Chomsky 1980(: 41-2), and Burton-Roberts and Carr 1999(: 383-4).

> The fact that a brown house has a brown exterior, not interior, appears to be a language universal... (1992b: 219; see also 1995a: 20)

That a brown house has a brown exterior rather than a brown interior, Chomsky argues, is an instance of the "intricate and highly-specialised" (1995a: 20) nature of the perspectives afforded by LEs for viewing the world -- the complexity of which

> ... poses problems of "poverty of stimulus" so extreme that *knowledge of language* in these regards... can only be assumed to be in substantial measure innately determined... (ibid, my italics)

It is not clear what exactly Chomsky means by "knowledge of language" in this case. The knowledge of "the fact that a brown house has a brown exterior, not interior"? There are two problems with this. Firstly, this knowledge seems to be the conscious, relational knowledge -- i.e. knowledge *of* something (i.e. the fact that a brown house has a brown exterior), rather than knowledge which is constitutive of something (i.e. I-language). Secondly, the fact that a brown house has a brown exterior cannot be a fact about the I-linguistic semantics of the LE 'a brown house'. *Having a brown exterior* obviously pertains to some E-physical brown house, or at least to the concept of some E-physical brown house. Neither of these can be in any direct relation to the I-linguistic semantics of the Chomskian LE 'a brown house', although one may be referred to, and the other expressed, through the *use* of the LE. So, what Chomsky calls a "language universal" appears to be a fact about the *use* of the LE 'a brown house', rather than a fact about the LE itself.

Returning to my main argument, my point is that Chomsky merely assumes, but never explains how, an LE like 'a brown house' is used literally to express *a house with a brown exterior*. Rather than giving an independent account of the semantics of the LE, Chomsky simply assumes that, since the LE 'a brown house' is generally used to express *a house*

with a brown exterior, the LE must provide "intricate" house-with-brown-exterior-orientated perspectives for viewing the world. In short, he simply projects aspects of the *use* of the LE onto the LE itself, contrary to his own assertions about the purely internalist and inclusive nature of I-language.

That Chomsky tends to attribute aspects of language use to LEs is clear from some of his remarks on the nature of I-linguistic semantics. For example, he maintains that LEs do not refer to things in the world, but can be used for that purpose *because their semantics provide perspectives for viewing those things*. He even goes as far as to make the following claim: that one of the *semantic properties* of a word is that the word "*can be used to refer to certain kinds of... things*" (1992b: 219, my italics). This suggests that, in virtue of providing such perspectives or having such semantic properties, the LEs of Chomsky's supposedly internalist theory are not independent entities, but merely instruments to be used -- or "*potentialities* of expression" (1995a: 26, my italics), as Chomsky himself puts it. These instruments or "potentialities" are therefore inextricably linked, and subordinate, to the performance systems whch employ them for the articulation, interpretation, expression of beliefs and desires etc.

In fact Chomsky himself states quite explicitly that I-language is subordinate to performance systems. Presumably LEs are used for expressing thoughts, or conveying meaningful communicative intentions, in virtue of having an LF interface, i.e. an interface with conceptual-intentional systems. But according to Chomsky, these LEs with their λs do not express any meaning. More importantly, Chomsky claims that *the I-language which generates LEs cannot be said to be "a language" unless it is embedded in the A-P and C-I performance systems*:

> It is only by virtue of its integration into [the A-P and C-I] performance systems that [I-language] qualifies as a language. Some other organism might, in principle, have the same I-language (brain state) as Peter, but embedded in performance systems that use it for locomotion. (1992b: 213)

It appears that this generative device which Chomsky calls the *I-language* cannot be said to be *anything at all* unless it is embedded in some sort of performance system which makes use of it in some way. This means that the semantics of the I-language is a semantics only in virtue of having the capacity to be used by certain performance systems. So it would seem that Chomsky's supposedly internalist LEs are not so much *used because they have a semantics*, as *endowed with a semantics in virtue of their capacity to be used -- and furthermore, used in a culturally, historically and completely accidentally determined manner*.

This problematic relation between I-language and performance systems can also be perceived in the technical detail of the Minimalist Program. It could be argued that the whole architecture of the I-language, as described in the Minimalist Program, is geared towards the condition of Full Interpretation (FI) at the interfaces with the performance systems. FI has to be satisfied by the PF and LF representations at the interfaces, in order for a derivation to converge. It follows from this that I-language and the LEs it generates are subject to *external factors* which have to do with extra-linguistic *performance*. To satisfy FI at LF, λ must (i) provide the C-I performance systems with "appropriate instructions"; and (ii) be assigned "uniform, language-independent interpretations" by the C-I systems (1995b: 194). This means that convergence at the interface, and consequently whether and what LEs emerge from C_{HL}, is at least partly decided by the *performance systems*.[21]

The role played by the performance systems in the generating of LEs comes across more clearly in the case of π and the PF interface. In 2.1.3 I noted that the nature of PF is such that the process of language use for humans must in some direct or indirect way involve the production and interpretation of E-physical utterance phenomena. According

[21] Chomsky himself admits that the principle of inclusiveness is violated at both PF (1995b: 229, 236) and LF (ibid: 227).

to Chomsky, this has to do with the "bare output conditions" imposed at the interface by "systems that make use of the information provided by C_{HL}": at PF, the information which constitutes π "has to be accommodated to the human sensory and motor apparatus" (1995b: 221). But Chomsky adds:

> Hence U[niversal] G[rammar] must provide for a phonological component that converts the objects generated by the language L to a form that these "external" systems can use: PF, we assume. If humans could communicate by telepathy, there would be no need for a phonological component, at least for the purposes of communication; and the same extends to the use of language generally. (1995b: 221)

What Chomsky seems to be saying here is this: it is only the non-telepathic human's need to communicate through the production and interpretation of E-physical utterances that demands a phonological component in the human I-language, and a PF representation in the LE. This means that the nature of I-language is profoundly influenced by the performance systems in which it happens to be embedded, and the way it happens to be used.

2.3 Chomsky on the "misuse of language"

The "misuse of language" is a topic which will be discussed in much greater detail in chapters 4 and 5. My aim here is to set out Chomsky's views, not so much on the phenomenon of the "misuse of language" itself, but on what he considers to be a confusion of the misuse of I-language with the misuse of certain non-linguistic entities. I intend to demonstrate that, without an independent account of how I-language constrains the use of I-language, Chomsky actually has no grounds for distinguishing between *the use/ misuse*

of an I-language which is not "designed for use" (1995b: 18), on the one hand, and *the use/ misuse of entities specifically constructed to be used by speaker-hearers*, on the other.

Chomsky's discussion of "misuse" is in response to Davidson 1986. Davidson argues that our ability to correctly interpret malapropisms, new words and "incomplete or grammatically garbled" utterances (1986: 437) indicates that there is

> ... no learnable common core of consistent behaviour, no shared grammar or rules, no portable interpreting machine set to grind out the meaning of an arbitrary utterance. (1986: 445)

Davidson points out that any speaker, "ingenious or ignorant" (ibid: 441), may convert any word or construction to a new use, or even invent completely new words. And yet speakers and interpreters "get away with it" all the time -- they may arrive at an occasion of utterance with different theories of determining utterance meaning, but

> ... the speaker is nevertheless understood; the interpreter adjusts his theory so that it yields the speaker's intended interpretation. (ibid: 440)

Thus Davidson concludes that, if language is regarded as "a clearly defined shared structure which language-users acquire and then apply to cases", then "there is no such thing as a language" (ibid: 446).

For Chomsky, however, language is *I-language*, upon which the production and interpretation of utterances, or the use or misuse of language, is supposed to have no bearing whatsoever. Chomsky describes three types of "misuse of language", all of which (he claims) "[play] no important role in the study of language, meaning, communication, or whatever" (1992a: 122). There is "misuse" in the "individual sense":

> Jones may speak in a way that is not in accord with his I-language, or may offer judgments inconsistent with his I-language... (ibid: 121)

When one speaks in a way that is not in accord with one's I-language, it appears that one tends to do so deliberately:

> He may choose to violate the rules [of his I-language], perhaps using the word "chair" to mean table in a code -- knowing that in his own language it means chair... (1995a: 36)

However, Jones may or may not be aware of the inconsistency of his judgement with his I-language when he "[misinterprets] an expression, in that his performance system yields an interpretation different from the one his internal language imposes" (ibid). Nevertheless, speaking, judging and interpreting are all instances of behaviour, and "much more than I-language is involved in behavior" (1992a: 121).

Chomsky also mentions "misuse of language" in the "community sense". Examples of this include Jones using the word 'disinterested' to mean *uninterested*, or his native dialect in a formal lecture (1992a: 121, 1995a: 36). Or

> ... Jones may try to adapt to the practice of some community for some reason, or perhaps for no reason at all, and may fail to do so, in which case people observing Jones may speak informally of a misuse of the language of this community. (1992a:121)

Chomsky claims that what Jones misuses or fails to conform to in these cases are *not in fact I-linguistic principles or I-linguistic expressions*, but "community norms" or "social practices" (ibid). These "community norms" are probably set up by certain authority figures, and observed by most of the members of the social group to which Jones belongs. As Chomsky points out (ibid), such "community norms" may be "of interest for the study of the sociology of group identification, authority structure, and the like".

The third and last sort of "misuse of language" which Chomsky discusses is misuse in the "expert sense". This "derives from Hilary Putnam's [(Putnam 1975a)] notion of "the division of linguistic labour" " (Chomsky 1992a: 121): one may be said to misuse language in the "expert sense" when one's application of terms like 'elm', 'acid' or 'mass' does not correspond to the way they are used by experts to whom one defers. But what is misused in the "expert sense" is (as in the case of the "community sense") *not* language as defined by Chomsky, but symbolic systems constructed for picking out objects in the world or expressing shared thoughts (1995a: 26). Such systems constitute "the Fregean ideal", as Chomsky calls it (ibid); and belong to naturalistic inquiry and perhaps a special "science-forming faculty", rather than to ordinary usage and the language faculty. Chomsky argues that the constructions of naturalistic inquiry and the science-forming faculty -- like 'mass' or 'momentum' in the physicist's sense, 'elm' in the botanist's sense, or 'acid' in the chemist's sense -- have origins and properties different from, say, 'house', and therefore "do not really belong to natural language" (ibid: 25). Instead of

> ... [entering] the lexicon by the same mechanisms of the language faculty that allows a child to pick up such words as "house" or "rise"... (ibid: 46)

naturalistic inquiry terms "abstract from the intricate properties of natural language expressions" (1992b: 209) and are "assigned meaning in a considered and determinate fashion" (ibid: 208). Concepts of common-sense understanding, on the other hand, "simply "grow in the mind", much in the way that the embryo grows into a person" (ibid).

Also, the meaning assigned to a naturalistic inquiry term consists of "semantic properties that may well not hold for natural language, such as reference" (ibid: 209). Chomsky maintains that naturalistic inquiry terms are

> ... divested of distorting residues of common-sense understanding and... assigned a relation to posited entities and a place in a matrix of principles. (ibid: 208)

In short, his claim is that naturalistic inquiry terms are constructed for the purpose of picking out things in the world -- their semantics based on Frege's *Bedeutung*, a technical relation between symbols and things (ibid: 226, 1995a: 26) -- while the semantics of natural language LEs is entirely intensional and internal (as outlined in section 2.2.2).

It would appear that at least two out of Chomsky's three types of "misuse of language" -- the "community" and "expert" senses -- are irrelevant to the study of language because, in either sense, it is not I-language which is misused. In other words, it is not LEs, but the constructions of social practice and naturalistic inquiry, which are (respectively) the objects of "community" and "expert" misuse. So it is only in the "individual" sense that the I-language is misused and, of course, *used*. This, then, is all that we may deduce about literalness within the Chomskian framework: the literal use of language is *to speak or judge in a way that is in accord with one's I-language*. According to Chomsky, Jones misuses his I-language when he uses 'table' to mean *chair*. It follows from this that Jones' use of his I-language is consistent with the I-language itself when he uses 'table' to mean *table* and (presumably) 'chair' to mean *chair*. (Or when he interprets 'table' as meaning *table* etc.) But not when he uses 'disinterested' to mean *disinterested* and 'uninterested' to mean *uninterested*, apparently, since the 'disinterested'-'uninterested' distinction is supposed to be a community, not an I-linguistic, norm. And not when he uses 'momentum' to mean whatever it is that physicists refer to when they use the term, because 'momentum' in that sense is not an LE, but a construction of naturalistic inquiry and the science-forming faculty.

Clearly this is not saying very much about either literalness or the use of I-language. But, vague and uninformative though it is, a problem nevertheless arises as soon as this formulation of "individual" language use is considered side by side with the examples of

adherence to community or scientific norms. The problem is that the distinction between these examples and examples of use in the "individual" sense may not be as well-defined as it should be.

To begin with, the use of 'table'/'chair' may not have any *more* to do with I-language than the use of 'disinterested'/'uninterested' or 'momentum'. According to Chomsky, the most obvious difference between the "individual" and the other two senses of use or misuse is that the use of 'table'/'chair' is a matter that is strictly between Jones and his I-language; while the use of 'disinterested'/'uninterested' and 'momentum' involves authority figures and experts, and their influence over groups of people. More specifically, Chomsky claims that the common-sense concepts which are expressed by the use of 'table'/'chair' grew naturally in Jones' mind "much in the way that the embryo grows into a person" (1992b: 208). In contrast, the 'disinterested'/'uninterested' conventions and the particular physicists' use of 'momentum' are set up, deliberately and without any individual I-language being taken into account, by authority figures and experts.

But apart from this, there is nothing to prevent Chomsky's example of "individual" use from changing places with the examples of the other two senses. It is perfectly possible that the use of 'table' to mean *table* is a community norm in some other community; just as it is possible for the use of 'disinterested' to mean *disinterested*, or 'momentum' to refer to some posited entity, to constitute behaviour that is in accord with someone's I-language. The question is, just how natural and individual is the process by which Jones learns to use his supposedly I-linguistic 'table' to mean *table*, compared to the process of learning the community norm of using 'disinterested' to mean *disinterested*? Surely Jones had to defer to other speakers in his environment about 'table', as he does to authority figures and experts about 'disinterested' and 'momentum'. From this point of view it does appear that "individual" use is not as I-linguistically determined, and as different from "community" and "expert" use, as Chomsky claims it is.

There is a further blurring of this distinction when one considers the possibility that the use of 'disinterested'/'uninterested' and 'momentum' may not have any *less* to do with I-language than the use of 'table'/'chair'. Chomsky himself suggests this, in two ways. Firstly: while he states that naturalistic inquiry terms "do not really belong to natural language" and "do not have the properties of natural language" (1995a: 26), Chomsky nevertheless concedes that we may articulate them "with the phonetics of our language", and "borrow constructions of our language in using them" (ibid); or construct them from "resources of the I-language (pronunciation, morphology, sentence structure, etc.)" (1992b: 226). The same may be said of terms and constructions set up by authority figures as "community norms".

Secondly, Chomsky appears to be claiming that changes in one's I-language are created when one adopts naturalistic inquiry terms like 'momentum', or community norms like the use of 'arthritis' to mean a pain in the thigh (1992a: 122). Constructions of naturalistic inquiry are "language-like accretions to the I-language"[22] (1992b: 226), while one's modification of one's use of, say, 'arthritis' may be a "marginal and rather arbitrary [variation] of I-language" (ibid: 228).

All this is worrying -- not just because of the implied transformation of what was originally excluded from natural language into "accretions to" or parts of I-language, but also because it seems to contradict the fundamental Chomskian idea that it is the I-language which determines language use, and not vice versa:

> The fact remains that Jones speaks and understands the way he does on the basis of the I-language he has acquired in the course of language growth; and if Jones does or does not follow what we choose, for some transient purpose, to call "community norms" or

---

[22] "Language-like accretions" is in fact an almost infinitely ambiguous and suggestive phrase. For one thing, what exactly does "language-like" mean: does "language" refer to *I-language*, and what is it to be "like" language? Another question is: if the accretions are merely "language-like" rather than "linguistic", then why are they accreting to the I-language?

> "social practice", it is on the basis of this internalized I-language (along with much else). (1992a:123)

For Chomsky, it is only in the "individual" sense that I-language is used. Thus the "individual" sense of use is distinguished from the "community" and "expert" varieties -- the constructions and processes of which, according to Chomsky, are irrelevant to the study of language precisely because they are not in any way determined by I-language. But without an independent account of I-linguistic semantics, and of how the "individual" use of an I-language is determined by the I-language, Chomsky has no way of showing that the "individual" use is in any way different from the other kinds of use. More specifically, he has no way of showing that the object of "individual" use is I-language, while the object of "community" or "expert" use is not I-language. Furthermore, his description of "individual" use as "speaking/ judging in a way that is in accord with one's I-language" is reduced to meaninglessness by his claim that the terms of community norms and naturalistic inquiry are constructed from elements of, and may eventually be incorporated into, I-language.

2.4 Conclusion

The notion of literalness -- of there being a fixed, direct way of expressing and recovering meaning -- appears to be closely related to two notions about language:

> (11) Language is a set of principles that leads us quickly and directly from vehicle to meaning.

(12) Language is autonomous -- it is crucial that the set of linguistic principles is in some way set apart from the complexities and irregularities of thought and behaviour which the phenomenon of metaphor in effect represents.

While not exactly in conflict, these two notions require a link. Because, strictly speaking, it is not language but people that express and recover meaning, it is necessary for us to have some sort of mental hold on language if it is to do what we expect of it. Hence the concept of language USE, according to which one *uses* language to express and recover meaning. In this way we end up with two loci of semantic content -- language, on the one hand, and the meaning which we sometimes express through the use of language, on the other.

Chomsky's intensionalist and internalist theory of language appears to be intended, and is generally assumed, to be the account *par excellence* of language as an autonomous set of principles. In fact it could be said that he goes one further in claiming that I-language is *not designed for use* (1995b: 18); and that its semantics is therefore not *semantic* in the sense of pertaining to the sort of propositional meaning which humans find significant. In other words, one of the main points of Chomsky's theory is this: even though it does happen to be used (in humans) to express and recover propositional meaning, I-language is not necessarily *for* the expressing and recovering of propositional meaning (or, for that matter, anything else).

This may be a perfectly logical progression from notion (12), i.e. the notion of the autonomy of language. However, it also exacerbates the problem of how an autonomous linguistic semantics may be harnessed for the characteristically human purpose of entertaining, expressing and recovering propositional meaning, by raising two additional questions:

(13) What exactly does an autonomous non-propositional semantics consist of?

(14) How is it accessed and converted into the propositional meaning which humans are interested in?

The problem is conceptual as well as empirical, and one of the arguments in this chapter is that it is not solved by any concept of "language use", or by the postulation of performance systems that interface with I-language. Another problem with the extreme, non-propositional austerity of Chomsky's version of linguistic semantics has to do with his shifting the entire burden of propositionality to the territory of language use -- this, together with the unanswered questions (13) and (14), creates a dangerous (conceptual/ empirical) gap between I-language and language use. I have argued that Chomsky ignores the problems mentioned above, and simply assumes that there is a particular use of an I-language which is "in accord with" that I-language. The result of this assumption is that, instead of giving an independent account of what an I-language and its semantics consists of, Chomsky ends up attributing to I-language some of the properties of the products of language use, thus contradicting the original, crucial notion of language's autonomy.

In 2.2 I pointed out that Chomsky's definition of I-language and I-linguistic semantics is undermined by the extent to which it is bound up with, and determined by, the performance systems which effect the use of I-language. If I-language is so closely related to the performance systems as to be subject to the influence of how it is used by humans, then clearly its semantics does not really count for anything -- the corollary being that we can bend I-language to our will, use any LE to express anything. These remarks of Chomsky's on language use seem to bear it out:

> ... person *X* uses expression *E* with its intrinsic semantic properties to talk about the world from certain intricate perspectives, focusing attention on particular aspects of it, under circumstances *C*, with the "locality of content" they induce... THE COMPONENTS OF *E* MAY HAVE NO INTRINSIC SEMANTIC RELATION AT ALL TO WHAT JONES IS REFERRING TO, as when he says the

> performance at Jordan Hall was remarkable, referring to Boston and his favourite string quartet. (1995a: 43-4, my caps)

This leads me back to the notion of literalness; and to suggest that literalness cannot in fact be defined in terms of a fixed and autonomous set of I-linguistic principles, for the lack of an intelligible concept of "language use" or whatever it is that is required for mediating between internalist linguistic semantics and cognisable, expressible literal meaning. So perhaps it is the other way round -- perhaps it is the notion of literalness which is the basis for theories of language like Chomsky's, and the root of all the problems which I have mentioned so far. The fact is that literalness is, for all its apparent stability and autonomy, as inseparable from thought and behaviour as metaphor is, and therefore no basis for an internalist linguistic theory. And yet the very concept of an internalist linguistics appears to have arisen from the notion of literalness. Consequently, *all* truly internalist theories of language may be doomed to circularity.

3

The distinction between linguistic and truth-theoretic semantics: literal meaning and relevance theory

3 Introduction

In the previous chapter I claimed that literal meaning has to do with the *use* of language, rather than with *language itself*. I argued that such a notion of literal meaning follows from Chomsky's internalist linguistics. In this chapter I argue that it also follows from the modular view of language that appears to underpin Sperber and Wilson's account of verbal communication. Sections 3.1 and 3.1.1 are concerned with some of the general theoretical background of Sperber and Wilson's (1986a/95) *relevance theory*. In 3.1.2 I set out Sperber and Wilson's respective definitions of their *principle of relevance* and what they call *ostensive-inferential communication*.

In chapter 2 I observed that Chomsky's linguistic semantics is constituted by LF representations -- "mental representations" which have no relation to, and are therefore neither true nor false of, things in the world. For Chomsky's linguistic expressions to express the sort of propositional and entertainable meanings which bear upon "human interests, intentions, goals and actions" -- to express a concept even as elementary as *nameable thing* -- their LF representations have to interface with the C-I performance systems, i.e. have to be *used* (by an agent).

Sperber and Wilson, whose relevance theory is the subject of this chapter, also subscribe to the notion of linguistic semantic representations as mental objects that "never surface to consciousness" (1986a/95: 193) -- and therefore as distinct from the sort of truth-theoretic propositional meaning which could be said to be literal (or non-literal). So within the framework of classical relevance theory, literalness (or non-literalness) does not

apply at the level of the linguistic expression itself (i.e. the linguistic semantics), but at the level of *what is (literally) said* through the uttering of the linguistic expression (i.e. the truth-theoretic semantics).

Thus both Chomsky and the relevance theorists regard language as an autonomous I-language or language module, setting its "semantics" apart from the "*real*", truth-theoretic semantics of what the *language user* expresses. However, neither Chomsky nor Sperber and Wilson seem able to avoid confusing or conflating one with the other. As I argued in chapter 2, Chomsky begins by positing a truly autonomous generative procedure, but ends up defining I-linguistic semantics in terms of how LF representations interface with the C-I systems (and, more generally, defining the I-language in terms of how it interfaces with the performance systems). Sperber and Wilson's problem, ironically, is that -- unlike Chomsky, who simply asserts that his I-linguistic semantics is independent of the performance systems in which it is (or can be) embedded -- they attempt to provide some sort of semantic grounds for their distinction between linguistic and truth-theoretic semantics. In sections 3.2-3.3.1 I argue that each of the two assumptions upon which the distinction appears to be based presupposes in its turn that linguistic and truth-theoretic semantics are related in such a way as to override the distinction which the relevance theorists seek to make.

The first of these two assumptions (discussed in sections 3.2-3.2.1) is that many (but, problematically, *not all*) linguistic expressions are non-truth-conditional and non-truth-evaluable in virtue of encoding "semantically incomplete" logical forms. However, Sperber and Wilson also claim that such logical forms undergo formal logical operations like truth-theoretic propositions do, and enter into logical relations with truth-theoretic propositions. What is more, it is essential that a linguistically encoded logical form has these features, in order that it may be "*developed*" into a truth-theoretic propositional form -- i.e. so that it may be *used* to express a truth-theoretic propositional form. The relevance-theoretic process of *logical form development* is outlined in the expository

section 3.2, while in section 3.2.1 I discuss and criticise Sperber and Wilson's idea that there are "semantically incomplete" logical forms which have to be "developed".

The second assumption which underlies the relevance-theoretic distinction between linguistic and truth-theoretic semantics (discussed in 3.3) is that there are some truth-evaluable, fully propositional forms which are nevertheless non-truth-theoretic in virtue of being "trivially true or blatantly false" (Carston 1998: 160) and/ or "too unspecified" (ibid: 164) to be something which "humans are interested in communicating" (Sperber and Wilson 1986a/95: 174). Examples of these include

(a) It will take some time to get there.

(b) I haven't eaten.

Thus the second of Sperber and Wilson's reasons for distinguishing linguistic semantics from real, truth-theoretic semantics has to do with their view of truth-theoretic propositions as propositions which are worth entertaining or communicating -- in short, propositions which are *relevant*. This, of course, presupposes that Sperber and Wilson's "non-truth-theoretic" propositions and truth-theoretic propositional forms are interchangeable to the extent that one can measure the former's lack of relevance against the relevance of the latter.

I conclude that Sperber and Wilson's attempt to distinguish linguistic semantics from truth-theoretic semantics -- to set language apart from what it is used for -- not only fails on its own terms, but also creates problems for the relevance framework. Within the wider context of this thesis, my conclusions suggest that, while the idea of linguistic expressions as vehicle-meaning pairs does presuppose some sort of distinction between the meaning of the "linguistic" vehicle, and the meaning which the vehicle is used to express, it also presupposes a notion of language use -- of "linguistic" vehicles as

specifically for being uttered -- which an absolute ontological distinction between linguistic and truth-theoretic semantics is simply not consistent with.

3.1 Relevance theory's Fodorean foundations

Before I go into the relevance theory account of communication and non-literalness, it is necessary to discuss the Fodorean framework assumed by Sperber and Wilson on certain fundamentals concerning thought and language.

Firstly, Sperber and Wilson's relevance theory is predicated on the assumption that cognition is computational. This view of cognition -- which they take very much for granted -- is set out in some detail, and argued for, in Fodor's 1975 *The language of thought*. (See also Fodor 1987.) One of its most important aspects has to do with the fact that computation presupposes a medium for computation, a means of representing what the computations compute: i.e. a representational system or language. Thus Fodor posits an internal language, a "language of thought", in which cognitive processes are carried out. Take perception, for example, which he defines as the formation and confirmation of hypotheses about what the world is like (1975: 44-9). On his computational-representational approach, this would involve the inference (computation) of conclusions (representations of hypotheses constructed and chosen) from premises (representations of perceptual information, as well as information stored in the memory).

Secondly, Sperber and Wilson also subscribe to Fodor's modularity thesis (Fodor 1983), which distinguishes between input systems (or modules) and central systems. According to Fodor, all of these systems are computational and representational, but in different ways. The input systems perform computations over perceptual representations, and each input system only processes information in a representational format peculiar to itself -- the visual module only processing visual information, the auditory module only

processing acoustic information, and so on. Fodor calls this property of input systems "domain specificity" --

> ... only a relatively restricted class of stimulations can throw the switch that turns [an input system] on... (1983: 49)

-- while Sperber and Wilson describe input processes as "local":

> ... a local process is either context-free or sensitive only to contextual information from some set domain... (1986a/95: 65)

All these domain specific input systems function to convert perceptual information, represented in their different ways, to the single format which constitutes the domain for the operation of the central processes. Fodor argues that it is the central systems that integrate what the input systems (and memory) deliver, and "use this information to constrain the computation of 'best hypotheses' about what the world is like" (1983: 104). This means that the central systems are what Fodor calls "domain inspecific" (ibid: 103), in the sense that they take into account and operate over all available information, whether the information is provided by the input systems, stored in the memory ("background knowledge" (ibid: 102)) or, possibly, innate. Or, to use Sperber and Wilson's term, the central systems are "global" as opposed to "local", in that they have "free access to conceptual memory" (1986a/95: 65).

Thirdly, Sperber and Wilson share the Fodorean view of natural language utterances as perceptual information:

> ... utterances (e.g., sentence tokens) are themselves objects to be perceptually identified, just as mountains, teacups, and four-alarm fires are. (Fodor 1983: 44)

To be more precise, they see utterances of linguistic expressions as perceptual information in a particular representational format, processable only by the appropriate input system. They regard this input system as a specifically linguistic module which converts utterance information into conceptual representations over which the central systems can operate.

Note that although he sometimes refers to input systems as *perceptual* systems (e.g. 1983: 43-4), Fodor's theory of perception has it that nothing is perceived without the participation of the central systems. He argues that input processing of information about some aspect of the world is necessary but not sufficient for that aspect of the world to be actually perceived. This is because:

> ... perception is a mechanism of belief fixation par excellence: the normal consequence of a perceptual transaction is the acquisition of a perceptual belief. (ibid: 40)

In terms of computations and representations, the fixation of perceptual belief involves the confirmation of hypotheses about what the world is like, or is said to be like. It is not possible for any single input system, with its own computational and representational system, to perform such a task, since *all* the input systems (not to mention memory) provide information about the world, and any of this information may be relevant to the construction of representations of how the world appears to be (ibid: 102-3). The "global" central systems, on the other hand, do have access to the output of all the input systems -- in fact, presumably, to every source of conceptual representations that the mind has. Thus it is the central systems that carry out inferential processes on representations, while the input systems merely provide some of these representations. Or, as Fodor puts it, it is the central systems which "people have in mind when they talk, pretheoretically, of such mental processes as *thought and problem-solving*" (ibid: 103, my italics). The input systems, on the other hand, merely function to "*represent the world*" so as to make it "*accessible to thought*" (ibid: 40).

All this has to be borne in mind when one considers the particular case of utterance interpretation. The language module is, of course, an essential part of the process. Fodor points out that

> Understanding a token sentence presumably involves assigning it a structural description, this being part and parcel of computing a token-to-type relation; and that is precisely the sort of function we should expect an input system to perform. (ibid: 44)

But if understanding an utterance is a "typical perceptual process" (ibid: 45), then obviously it cannot consist of input computations alone. Or, to put it another way, utterance interpretation cannot simply be the matching of physical forms to concepts, or the conversion of particular types of sensory data (i.e. utterance phenomena, like acoustic or orthographic objects) into conceptual representations upon which central processes may operate. Following Fodor's notion of what perception involves, it is clear that an utterance is understood only when the conceptual output of the language module *is* operated upon by the central systems.

To summarise the main Fodorean ideas which relevance theory assumes:

> (1) Cognitive processes are computational, and have representational systems in which to compute.
>
> (2) There is a distinction between (A) systems whose function is to provide representations of information about the world (the INPUT SYSTEMS), and (B) systems which perform inferential processes over those representations (the CENTRAL SYSTEMS).
>
> (3) One of the input systems is a language module, which -- together with the central systems -- is essential for, and specific to, utterance interpretation.

(4) The distinction between the language module and central systems is coextensive with the distinction between what Sperber and Wilson call (A') "NATURAL" (or EXTERNAL) LANGUAGES, the representations of which form the domain for the operation of the language module; and (B') an INTERNAL LANGUAGE (or language of thought), the medium for central computations.

In conclusion, I must stress that it is Sperber and Wilson's *"natural" or external languages* – representational formats over which the linguistic input system operates -- that I am talking about when I use terms such as "linguistic", "linguistic expressions", "sentences", "linguistic semantic representations" and (obviously) "language module" in this chapter. When I discuss *internal language* – the medium for central computations -- I will refer to it as such (e.g. "formulae of the internal language").

3.1.1 Relevance theory's departures from Fodor

That the language module on its own cannot effect utterance interpretation is one of relevance theory's starting points. Sperber and Wilson certainly share Fodor's view of utterance interpretation as a largely *central* and therefore inferential process involving the construction and confirmation of hypotheses about the speaker's communicative intentions (Fodor 1975: 108, and fn7). From this they also conclude, explicitly, that linguistic expressions (LEs) and linguistic meanings in themselves -- without any interaction with the central systems, without any interaction with thought, i.e. *without being used* -- are of no immediate cognitive significance to utterance producer or interpreter:

> One entertains thoughts; one does not entertain semantic representations of sentences. Semantic representations of sentences are mental objects that never surface to consciousness. If they did, they would seem entirely uninteresting (except, of course, to semanticists). (1986a/95: 193)[1]

Or, as Carston (1998: 103) puts it, "We do not communicate logical forms (though we *do* communicate via logical forms)".

These remarks seem consistent with Fodor's conception of the input and central systems' respective roles in utterance interpretation, but in fact this is where Sperber and Wilson begin to diverge from Fodor. To begin with, the language module occupies a far more important place in Fodor's theory than in Sperber and Wilson's. For Fodor, a "natural" or external language is "a system of conventions for the expression of communicative intentions" (1975: 106), and

> Verbal communication is possible because the speaker and hearer both know what the conventions are and how to use them: What the speaker knows allows him to pick a value of [acoustic object] *U* which encodes a given value of *M* [what the speaker intends to communicate], and what the hearer knows allows him to pick the value of *M* which is encoded by a given value of *U*. (ibid: 108)

The knowledge of such linguistic conventions may be used to

> ... [effect] a certain correspondence between the mental states of speaker and hearer: The speaker is enabled to construct utterances which *do* express the messages that he intends them to express; the hearer is enabled to construe the communicative intentions of the speaker. (ibid)

---

[1] This last sentence is somewhat at odds with the point Sperber and Wilson are making -- that linguistic semantic representations are not what we communicate or entertain unless they have been incorporated with thought in some way . Within the relevance framework it is true that semantic representations of LEs "never surface to consciousness". But if they did, they *would* be interesting, and not just to semanticists -- they would be interesting simply in virtue of having surfaced to consciousness. They would not surface to consciousness if they were not being used, or thought of; and if they were being used or thought of they must be interesting.

In other words, Fodor assumes that there is a direct and systematic correspondence between "messages" and linguistically encoded semantic representations; and, therefore, that the recovery of "messages" is simply a matter of decoding (even though the actual *perception* of linguistically encoded semantic representations as "messages" is brought about only by the central processes). So it seems that this Fodor 1975 account of verbal communication is based on what Sperber and Wilson call the *code model* of communication (1986a/95: 2).

For Sperber and Wilson, on the other hand, there is no such concurrence between what the utterer of an LE intends to convey, and the semantic content of that LE. They argue that what a speaker communicates when he utters an LE is much more, and much more complex, than the semantic content of that LE. They also argue that this additional non-linguistic content has to be inferentially recovered. To observe that Sperber and Wilson differ from Fodor in this respect may seem somewhat trivial -- after all, there is no doubt that *pragmatic inference and implication* features prominently in relevance theory, while it is not clear if Fodor 1975 takes pragmatics into consideration at all. Thus one might argue that the difference is simply down to the absence of a semantic-pragmatic distinction, and Fodor's (1975) situating of all verbally communicated meaning in the semantics of "natural" or external languages.

However, there is at least one point upon which Sperber and Wilson's account of utterance interpretation uncontrovertibly diverges from Fodor's. They attribute a property (if one can call it a property) to linguistic meaning which practically guarantees the involvement of inferential processes, even if the central systems were not already deemed essential for utterance interpretation: they claim that LEs either are, or often are, SEMANTICALLY INCOMPLETE. According to Sperber and Wilson, a semantic representation of an LE is a *linguistically encoded logical form* -- a logical form being

> ... a well formed formula, a structured set of constituents, which undergoes formal logical operations determined by its stucture. (1986a/95: 72)

-- and that most of these linguistically encoded logical forms are "*semantically incomplete*" and therefore *non-propositional* and *incapable of being true or false* (ibid):

> ... the sense of a sentence is often an *incomplete* logical form. (ibid: 73, my italics)
>
> Semantic representations [i.e. linguistic ones] are incomplete logical forms, i.e. at best fragmentary representations of thoughts. We have argued that they are incomplete in more than one way: not just because they contain indeterminate referring expressions such as pronouns, but also because they contain underdefined contituents such as 'too', 'some time', or the genitive. (ibid: 193)

For example, they maintain that the pronouns in the sentence

(5) She carried it in her hand.

"do not correspond to definite concepts, but merely mark an unoccupied place where a concept might go" (1986a/95: 72-3). So the logical form which (5) encodes cannot be used to represent a state of affairs in a possible or actual world, and hence cannot be true or false.

Note that Sperber and Wilson seem to be simultaneously claiming that ALL linguistic semantic representations are "incomplete logical forms" (1986a/ 95: 193), and that linguistic semantic representations are "often" (ibid: 73) but NOT ALWAYS incomplete logical forms. It is clear that Sperber and Wilson do accept the existence of semantically complete, fully propositional LEs, for they state that

(5a) No one ever carried anything.

"is, or can be understood as, propositional" (ibid:73). But the words "*can be understood as [propositional]*" suggest that (5a) can also be "*understood as NON-propositional*". Those words also suggest that, for Sperber and Wilson, the notion of "what is propositional" is inseparable from the notion of "the proposition understood to have been expressed". Furthermore, if some linguistically encoded logical forms are or can be propositional, then surely all linguistically encoded logical forms are propositional, simply because those that are will be entailed by those that are supposed not to be.

I will be returning to this problematic distinction between the "semantically incomplete" and the fully propositional in section 3.2.1. Here, my main purpose is to make the following point: Sperber and Wilson agree with Fodor that verbal communication involves both the language module and the central systems, and that the language module decodes linguistically encoded semantic representations, but have a different view of the central systems' role in utterance interpretation. For Fodor this appears to be confined to the confirmation of perceptual belief, leaving the burden of propositionality to be carried by the language module. Sperber and Wilson, however, claim that the central systems also contribute -- and *must* contribute -- to what the speaker conveys.

These opposing views are based on a disagreement concerning the role -- or rather, roles -- of language in human thought and behaviour. But first, consider what the relevance theorists and Fodor *do* agree about. Like Fodor, Sperber and Wilson claim the following:

(6) A language is "a grammar-governed representational system" or "a set of semantically interpreted well-formed formulas" (Sperber and Wilson 1986a/ 95: 173).

(7) The formulas of languages are semantically interpreted in more ways than one (Fodor 1975: 65-6, 73) -- they may be "put into systematic correspondence... with the formulas of another language, with states of the user of the language, or with possible states of the world" (Sperber and Wilson 1986a/95: 173).

(8) How a language is semantically interpreted depends on what that language is for.

(9) Language is a medium for *cognition* (hence *internal language*) as well as communication --

> Language is an essential tool for the processing and memorising of information. Any organism or device with a memory must be able to represent past states of the world or of itself. Any organism or device with the ability to draw inferences must have a representational system whose formulas stand in both syntactic and semantic relations to each other. (Sperber and Wilson 1986a/95: 173)

-- and especially for the cognitive processes of communicating devices --

> Two devices capable of communicating with each other must also be capable of internally representing the information communicated, and must therefore have an internal language. (ibid: 174)

But, unlike Fodor, Sperber and Wilson stress that although *language* as defined in (6) is absolutely necessary for cognition, it is

> ... not a necessary medium for communication: non-coded communication exists. (ibid: 174)

They point out that there is another model of communication, the inferential model, according to which

> ... communication is achieved by the communicator providing evidence of her intentions and the audience inferring her intentions from the evidence. (ibid: 24)

They argue that purely inferential communication exists (ibid: 26) -- it is possible, for example, to communicate something through (or infer someone's communicative intentions from) movements of some part of the face or body, those movements being "evidence" of the communicator's intentions. More importantly, the output of linguistic decoding is also regarded by Sperber and Wilson as constituting evidence from which the audience infers the intentions of the person who provided the linguistic input (ibid: 27). It follows from this view of linguistic decoding (which in itself is perfectly consistent with Fodor's claim that the central processes are a necessary part of utterance interpretation), *together with the possibility of purely inferential communication*, that coded communication via linguistic encoding and decoding is subservient to inferential communication. In other words, if

> (10) inferential communication is possible without the involvement of the language module, and
>
> (11) the language module's sole purpose is to provide evidence for inferential communication --

then, as a means of communication, linguistic encoding and decoding must be secondary to the central inferential mechanisms.

And it is on the basis of (10) and (11) that Sperber and Wilson maintain: LEs need not -- and therefore *do not* -- encode everything that a speaker might want to communicate, since the interpretation process will invariably be completed (inferentially) by the central systems. Thus they argue that the semantics of "natural"/ external languages

"might be too weak to encode all humanly thinkable thoughts" (1986a/95: 193), in virtue of the secondary nature of their role in communication. Or, as Carston puts it,

> ... the language system does not have the resources to encode the propositions speakers succeed in expressing, and what I am suggesting here is that there is a very good reason for this: that sort of expressive power is redundant. A powerful 'mind-reading' capacity is employed in the interpretation of human behaviour quite generally... (1998: 55)

There is an even more fundamental level at which Sperber and Wilson's departures from Fodor may be considered. To Sperber and Wilson,

> The fact that the semantic representations of natural-language expressions are merely tools for inferential communication suggests that inferential communication had to exist before external languages developed: human external languages are of adaptive value only for a species already deeply involved in inferential communication. (1986a/95: 176)

Carston also observes that

> The wide application of this capacity [for inferential communication] in human cognitive activity and its presence in rudimentary form in apes, who lack a linguistic system, make it reasonable to suppose that the linguistic code evolved later than, or perhaps in step with, the capacity to attribute mental states. (1998: 55)

In contrast, Fodor thinks that the view of the central systems as "philogenetically prior" to the input systems (including, presumably, the linguistic one) is "[a] fairly dubious evolutionary [assumption]" (1983: 43), suggesting instead that input systems, "with their relatively rigid domain specificity and automaticity of function", could in fact be the "aboriginal prototypes" of inference-making central systems (ibid).

I have demonstrated that, for Sperber and Wilson, communication is predominantly inferential, and linguistic decoding very much a subsidiary mechanism. Hence they argue that there is a division of interpretational labour between the language module (decoding) and the central systems (inference) during verbal communication.

One of the consequences of these views is that LEs do not need to -- and therefore do not -- encode everything that the speaker might intend to communicate. This is clearly at odds with Katz's stronger interpretation of his "principle of effability" (1981: 226):

> Katz's principle of effability: Every proposition (thought) is expressible by some sentence in every natural language.
> Katz's stronger interpretation of the effability principle: Every thought is encoded by a sense of some sentence. (in Sperber and Wilson's words, 1986a/95: 191)

For Sperber and Wilson, not every thought is encoded by a "natural"/ external language sentence, because the nature of "natural" languages is such that they cannot encode certain aspects of thought, and the nature of verbal communication is such that they do not need to.[2] Reference, for example, is regarded by Sperber and Wilson as obviously *in*effable:

> ... two people may be able to think *of* the same man... without being able to think exactly the same thought, because they might not individuate him in exactly the same way. Similarly, by saying 'He has gone' I may induce in you a thought which is similar to mine in that it predicates the same thing (that he is gone) of the same individual, but which differs from mine in the way you fix the reference of 'He'. (1986a/95: 192-3)

---

[2] I am assuming that Katz agrees with Sperber and Wilson on what "thoughts" are -- i.e. conceptual representations (i.e. propositions in the internal language format).

More generally, Recanati (1987, 1995; cited in Carston 1998: 37-8) suggests that reference is always determined by context. In Recanati 1993, he argues that the linguistic semantic content of a proper name is not so much a meaning, as a sort of procedure or constraint specifying that the proper name refers to its bearer -- the identity of the bearer being, of course, a matter of context. He also proposes (in Recanati 1987) that the reference of a definite description always depends on the "domain of discourse", defined as "that with respect to which the speaker presents his or her utterance as true" (Recanati 1987: 62).

Thus Sperber and Wilson conclude:

> It seems to us neither paradoxical nor counterintuitive to say that there are thoughts that we cannot exactly share, and that communication can be successful without resulting in an exact duplication of thoughts in communicator and audience. We see communication as a matter of enlarging mutual cognitive environments, not of duplicating thoughts. (1986a/95: 193)

Carston expresses similar sentiments when she suggests that

> A speaker's choice of linguistic form takes account of the hearer's immediately accessible assumptions, encoding just what seems to be necessary to direct the hearer's inferential processes to the intended interpretation... Verbal communication, on this view, is not a means of thought duplication; the thought(s) that the speaker seeks to communicate are seldom, if ever, perfectly replicated in the mind of the audience... (1998:18)

3.1.2 Ostensive-inferential communication and the principle of relevance

As I have mentioned in the previous sections, Sperber and Wilson claim that there are two modes of communication; and that one, coded communication, is subservient to the other,

inferential communication. Coded communication involves, of course, a *code*: a system of message-signal pairs which enables information-processing devices to communicate by decoding signals into which messages are encoded. According to Sperber and Wilson, the Morse code is an example of a simple code -- "a straightforward list of message-signal pairs" (ibid: 4) -- while an example of a more complex code is the grammar of a language, which "pairs phonetic and semantic representations of sentences" (ibid: 9). Language thus defined is therefore a medium for coded communication. However, Sperber and Wilson argue that this type of coded communication is an essential component of, but *does not constitute*, verbal communication. This is because they regard verbal communication as a particular kind of *ostensive-inferential communication* that is strengthened by coded communication via "natural"/ external languages (1986a/95: 63).

Before a formal definition of ostensive-inferential communication can be given, the notions of (A) MANIFESTNESS, and of the communicator's (B) INFORMATIVE and (C) COMMUNICATIVE INTENTIONS, have to be introduced:

(A) An assumption is *manifest* to an individual at a given time "if and only if he is capable at that time of representing it mentally and accepting its representation as true or probably true" (ibid: 39). Thus to make an assumption manifest to someone is to present it as true or probably true, or to make it perceptible or inferable.

(B) A communicator's *informative intention* is to make manifest or more manifest to the audience a set of assumptions {*I*}(ibid: 58).

(C) A communicator's *communicative intention* is to make mutually manifest to audience and comunicator the latter's informative intention (ibid: 61).

This is Sperber and Wilson's definition of ostensive-inferential communication:

> *Ostensive-inferential communication*: the communicator produces a stimulus which makes it mutually manifest to communicator and audience that the communicator intends, by means of this stimulus, to make manifest or more manifest to the audience a set of assumptions {*I*}. (ibid: 63)

For a stimulus -- an ostensive stimulus -- to evince the communicator's intention to make manifest or more manifest { *I* } (i.e. to evince the communicator's *informative intention*), it has to (a) attract the audience's attention, and (b) focus the audience's attention on the communicator's *communicative intention* (ibid: 153). An example -- in fact the example *par excellence* -- of how (a) is achieved is the use of language in verbal communication. As Sperber and Wilson point out, "spoken utterances in one's own native language *automatically impinge on the attention*" (ibid, my italics).

As for (b): a stimulus may focus the audience's attention on the communicative intention in virtue of being RELEVANT if and only if the audience takes it to have been produced with a communicative intention -- i.e. if and only it is taken to be an ostensive stimulus. A phenomenon is *relevant* if it gives rise to large and numerous CONTEXTUAL EFFECTS when it is processed with minimum effort. Contextual effects arise from the interaction between new information conveyed by some phenomenon and the old information of a particular context. They come in three categories: the addition of new assumptions; the providing of evidence for, and therefore strengthening of, old assumptions; and the providing of evidence against old assumptions, leading to their abandonment. The relevance of a phenomenon is proportionate to its contextual effects and processing costs -- the stronger and more numerous the contextual effects, the smaller the processing costs, the more relevant the phenomenon.

When the communicator's informative intention has been made manifest, the addressee has to "construct possible interpretive hypotheses about the contents of {*I*}, and to choose the right one" (ibid: 165). Sperber and Wilson's PRINCIPLE OF RELEVANCE states that

> Every act of ostensive communication communicates the presumption of its own optimal relevance. (ibid: 158).

So one of the members of {*I*} must be the PRESUMPTION OF OPTIMAL RELEVANCE:

> *Presumption of optimal relevance*
> (a) The ostensive stimulus is relevant enough for it to be worth the addressee's effort to process it.
> (b) The ostensive stimulus is the most relevant one compatible with the communicator's abilities and preferences. (1995: 270)[3]

In other words, every act of ostensive communication carries a guarantee of its own optimal relevance. Hence the addressee will choose the hypothesis which he has reason to think will confirm the presumption of optimal relevance. This hypothesis should be the *first* accessible hypothesis consistent with the principle of relevance. According to the second part of the presumption of relevance, the communicator would not have used a stimulus which sends the addressee on an effort-wasting pursuit of more than one hypothesis (1986a/95: 168).

3.2 Sperber and Wilson's account of verbal communication: logical form development

In this section, Sperber and Wilson's account of verbal communication is described in terms of an "assumption construction" process which they refer to as LOGICAL FORM DEVELOPMENT.

---

[3] This "presumption of optimal relevance" is the revised version given in the "Postface" of the second edition of Sperber and Wilson's *Relevance: communication and cognition*.

Sperber and Wilson's example (1986a/95: 176-82) has Mary uttering the complex sound

(1) [Itlgetk∂Uld]

in the presence of Peter. Being "a perceptible modification of the physical environment" (ibid: 176), this utterance makes manifest to Peter a set of assumptions {*A*}. The members of {*A*} include

(2)(a) Someone has made a sound.

(b) Mary is at home.

But since (1) is, specifically, a "*linguistic stimulus*", it also makes manifest the assumption that

(3) Mary has uttered the *sentence* 'It will get cold.'

As mentioned above, Sperber and Wilson claim that utterances like (1) "automatically impinge on the attention" (ibid: 153). In terms of Fodor 1983, this is because linguistic decoding is performed by an input system, and, like all other input processes, performed automatically on the system's being triggered by some "linguistic stimulus". Hence a semantic representation will be automatically assigned to (1), making manifest assumption (3) in the process.

Verbal communication proper begins "when an utterance... is manifestly chosen by the speaker for its semantic properties" (1986a/95: 178). Sperber and Wilson point out that we usually *do* assume that utterances are used for communication in virtue of their linguistic semantic properties -- the reason being that the first accessible interpretation of

an utterance that is consistent with the principle of relevance tends to be based on that utterance's automatically (and therefore very easily) recoverable linguistic semantic representation (ibid). So Mary's utterance (1) would also make manifest the assumption

(4) Mary has said to Peter 'It will get cold.'

and, "since saying something to someone is a case of ostensive communication" (ibid), the assumption

(5) There is a set of assumptions { *I* } which Mary intends to make manifest to Peter by saying to him 'It will get cold.'

as well. Having made (5) manifest, Mary would have achieved her *communicative intention* -- i.e. she has made it mutually manifest to herself and Peter that she intends to make {*I*} manifest to Peter. And if Peter "both understands her and trusts her enough" (ibid: 179), then she would also have achieved her *informative intention* -- i.e. her intention to make {*I*} manifest to Peter. In that case, {*I*} would be a subset of {*A*}, and Peter's task would be to decide which members of {*A*} were also members of {*I*}. In short, Peter's task is to decide which of the assumptions *made manifest* by Mary's utterance are also *communicated*, or *intended to have been made manifest*, by the utterance -- i.e. to identify the EXPLICATURES and IMPLICATURES of the utterance. Sperber and Wilson's definition of explicitness will be given at the end of this section. At this point it is sufficient to state that the main difference between explicatures and implicatures is that an explicature has some linguistically encoded semantic representation as a "sub-part", while an implicature does not.

The rest of this section deals with the progress from *linguistic semantic representation* via *propositional form* to *explicature* by way of logical form development.

After it has been made manifest to him that

(5) There is a set of assumptions { *I* } which Mary intends to make manifest to Peter by saying to him 'It will get cold.'

Peter clearly has to make something *propositional* out of the supposedly *non-propositional* logical form which he had recovered through linguistic decoding, before he can even begin to identify the *fully propositional* members of {I} which have that logical form as a sub-part (i.e. the explicatures).

According to Sperber and Wilson, the task of assigning a unique PROPOSITIONAL FORM to an utterance is an inferential one, and involves disambiguation, reference assignment and general enrichment of the utterance's linguistically encoded logical form (1986a/95: 179, 185-91). In the case of Mary's utterance, Peter has to decide whether 'cold' means *experiencing cold* or *inducing cold*, assign a referent to 'it', and enrich the meaning of "vague items" like 'will' (bid: 179). If, for example, he decides that 'it' refers to the dinner and 'cold' means *inducing cold*, and adds *very soon* to the meaning of 'will', then he would take the propositional form of Mary's utterance to be

(6) The dinner will get cold very soon.

It should be stressed that (6) is merely what Peter assumes to be the propositional form of Mary's utterance, and may or may not be a member of the set of assumptions {I} *communicated* by the utterance. In other words, we should remember that Peter would not know whether or not (6) is the explicature of the utterance -- not until he has identified the mood in which (6) is expressed, and, more importantly, the propositional

attitude which Mary intended to communicate (ibid: 180). Sperber and Wilson claim that mood is linguistically determined, and underdetermines propositional attitude.

According to Sperber and Wilson (ibid: 180-1), the propositional form (6) is an explicature only if the propositional attitude is one of *assertion*. In other words, Mary *communicates* (6) if and only if Peter assumes that she has *asserted* it. Sperber and Wilson argue that it is not enough for him to recover the assumption

(7) Mary has said that the dinner will get cold very soon.

-- because

> ... in the weak sense of 'saying that' which corresponds to the declarative mood, one can say that *P* without communicating that one believes that *P*. For example, in saying that the dinner will get cold very soon, Mary might be speaking metaphorically or ironically, in which case she would not communicate that she believes that the dinner will get cold very soon. (ibid: 180)

Nor do they think it is enough for Peter to assume that

(8) Mary believes that the dinner will get cold very soon

-- because

> A speaker who communicates that she believes that P does not automatically communicate that P. For instance, suppose it is mutually manifest that Peter believes that the dinner will stay hot for as long as it takes him to finish what he is doing, and that he has no reason to trust Mary's opinion more than his own. Then Mary could not have intended her utterance to achieve relevance by making manifest to Peter that the dinner would get cold very soon, but only by making manifest that she believes it will. (ibid: 181)

It is only when he decides that Mary is in fact *asserting* that the dinner will get cold very soon, that Peter would assume that she intended (6), as well as (7)-(8), to be made manifest by her utterance. Sperber and Wilson refer to such an utterance -- one which communicates its propositional form; or rather, one whose propositional form is its explicature -- as an ORDINARY ASSERTION. They also describe (6)-(8) as having "as sub-parts one of the logical forms encoded by the utterance" (ibid: 181). More specifically, each of (6)-(8) is a DEVELOPMENT OF THE LOGICAL FORM encoded by Mary's utterance -- having been

> ... constructed inferentially, by using contextual information to complete and enrich this logical form into a propositional form, which is then optionally embedded into an assumption schema typically expressing an attitude to it' (ibid).

This notion of logical form development is significant in that it is an integral part of Sperber and Wilson's definition of EXPLICITNESS:

> *Explicitness*
> An assumption communicated by an utterance *U* is *explicit* if and only if it is a development of a logical form encoded by *U*. (ibid: 182)

(6)-(8), therefore, constitute the explicit content of Mary's utterance. In short, (6)-(8) are the EXPLICATURES of Mary's utterance. And what Sperber and Wilson would call an *ordinary assertion* is an utterance whose propositional form constitutes one of its explicatures. This, then, is the difference between Sperber and Wilson's explicatures and implicatures: both explicature and implicature are assumptions communicated by an utterance, but only the explicature is a *development of the logical form* encoded by the utterance. An explicature usually does have some non-linguistic content -- logical form

development is an *inferential* process -- but the difference lies in the fact that explicatures have linguistically encoded logical forms as "sub-parts", while the derivation of implicatures is *entirely* inferential.

Here is a summary of the respective places of (linguistically encoded) logical form, propositional form and explicature within Sperber and Wilson's account of verbal communication.

The hearer who has recovered only the linguistically encoded LOGICAL FORM of the sentence uttered by the speaker knows that:

> The speaker has uttered "S" (S: the sentence which encodes the logical form)

-- whether or not he assumes that the speaker is communicating something by uttering the sentence. Even if he did assume that something was communicated, he would not have made any assumptions about what it is.

Nor would he have made any assumptions about what was *said* -- not, at least, until he has constructed the PROPOSITIONAL FORM of the utterance from the logical form. The hearer who has identified the propositional form P assumes that

> The speaker has said that P.

Such an assumption is an EXPLICATURE in itself, since the speaker who intended to communicate something by saying that P must at least have communicated that he has said that P. But it is only when the hearer assumes that P has been *asserted* by the speaker, that he would be able to identify P as an explicature -- i.e. assume that the speaker has communicated P.

Therefore --

LOGICAL FORM: what is *encoded* by the speaker's utterance.

PROPOSITIONAL FORM: what is *said* by the speaker.

EXPLICATURE: what is *communicated, or intended to be made manifest*, by the speaker.

In view of Sperber and Wilson's account of non-literalness (which will be discussed in 3.3), it is necessary to point out that *there appears to be two ways of developing a given logical form.*

According to Sperber and Wilson, logical form development involves the completion and enrichment of a logical form with contextual information, and its optional embedding into an assumption schema expressing an attitude to it. They also refer to a logical form as being a "SUB-PART" of any assumption derived through logical form development. What I want to point out is that there appears to be two different kinds of logical form development, *because it follows from Sperber and Wilson's definition that there are two different senses of "sub-part"*. Both kinds of logical form development involve the completion and enrichment of the logical form into a propositional form. The difference lies in the fact that the propositional form is embedded into an assumption schema expressing an attitude to it, during what I shall refer to as COMMUNICATIVE DEVELOPMENT, but not during the other kind of logical form development which I shall call LOGICAL DEVELOPMENT.

Assuming, as Sperber and Wilson do, that logical forms are "amenable to logical processing", having the logical form as a "sub-part" would be equivalent to having it as a *logical implication* in the case of logical development, but not in the case of communicative development (hence "*logical*" and "*communicative*" respectively).

For example, the propositional forms (7) and (8)

(7) Mary has said that the dinner will get cold very soon.

(8) Mary believes that the dinner will get cold very soon.

arise from the *communicative development* of the logical form linguistically encoded by Mary's utterance "It will get cold", while

(6) The dinner will get cold very soon.

is a *logical development* of the logical form.

It is curious that assumptions about propositional attitude should be "*optional*", for the choice of one kind of logical form development over the other clearly plays a crucial role in the identification of the explicature (and implicatures) of an utterance. For example, an utterance is an ordinary assertion only if the logical form it encodes goes through *logical development*, and not *communicative development* -- or at least goes through logical development independently of going through communicative development. Put another way, the propositional form of an utterance is an explicature if and only if the *propositional form alone* -- without any attitudinal embellishment, as it were, and therefore logically implying the logical form -- is presented as true, made manifest, communicated.

A non-literal utterance, on the other hand, is one whose propositional form P is not *explicated*, because the hearer assumes that the speaker has merely *said that* P, or said that he *believed that* P, or said P *metaphorically*, etc. In other words, a non-literal utterance is one whose logical form undergoes *communicative development*. Sperber and Wilson claim that the "only obvious explicature" (1986a/95: 225) of a metaphorical utterance like

(9) This room is a pigsty.

is

(9a) The speaker is saying that this room is a pigsty.

In terms of Sperber and Wilson's definition of logical form development, there is only one way of developing a metaphorical utterance's logical form into something fit for presenting as true (i.e. fit for explicating), and that is to integrate the inferentially enriched, fully propositional version of the logical form -- the utterance's propositional form P -- into the appropriate assumption schema to give

(9b) The speaker said that P.

(9a), of course, does not logically imply the logical form or the propositional form of the metaphorical utterance (9). The most important part -- in fact the most relevant part -- of what (9) communicates is communicated implicitly -- i.e. in the form of implicatures which, unlike explicatures, are not developed from the linguistically encoded logical form. Therefore, since the propositional form of a metaphorical utterance is not explicated, and since the propositional form of an utterance may not be what the speaker believes, or intends to communicate, but is -- as (9a) shows -- definitely what the speaker has "SAID", we can only conclude that a metaphorical utterance, for Sperber and Wilson, involves *saying something which one has no intention of communicating*. What is more, since the only explicature of a metaphorical utterance with propositional form P is (9b), the person who makes that utterance is in effect explicitly communicating that he is saying what he has no intention of communicating. In section 3.3 I will be arguing that Sperber and

Wilson's conception of propositional forms and their truth-theoretic semantics creates problems for the relevance-theoretic account of utterance interpretation, especially in the case of non-literal utterances which explicate, not their propositional forms, but their communicatively developed logical forms.

3.2.1 "Semantically incomplete" logical forms and the distinction between linguistic and truth-theoretic semantics

Following on from my description of the relevance-theoretic process of logical form development in the previous section, this section raises questions about the nature of (a) Sperber and Wilson's *linguistically encoded logical forms*, and of (b) the relation between these logical forms and what they are developed into. In the process, we shall see how certain assumptions about (a) and (b) may have motivated and influenced Kempson's (1986) and Carston's (1988) distinction between LINGUISTIC and TRUTH-THEORETIC SEMANTICS.

As mentioned above, Sperber and Wilson maintain that most sentences have semantic representations that are *semantically incomplete logical forms* (hence the necessity of *developing* them into fully propositional forms for communicative purposes). They claim that a logical form has to be "semantically complete" to be capable of being true or false:

> ... that is, it must represent a state of affairs, in a possible or actual world, whose existence would make it true. (1986a/ 95: 72)

Thus Sperber and Wilson regard linguistic semantics as non-truth-conditional and non-truth-evaluable. However, they also argue that a conceptual representation *need not be "capable of being true or false" for it to be "amenable to logical processing"* -- it only

needs to be "well formed" (ibid, my italics). For example, Sperber and Wilson claim that (10) --

> (10) She carried it in her hand.

-- has a semantically incomplete logical form which can be neither true nor false. Nevertheless:

> In spite of its non-propositionality, [(10)] has obvious logical properties. For instance, it implies [(10a)], which is equally non-propositional, and it contradicts [(10b)], which is, or can be understood as, propositional:
> [(10a)] She held something in her hand.
> [(10b)] No one ever carried anything. (ibid: 73)

The problem is that it is not at all clear that logical forms like (10)'s are really non-truth-conditional and non-truth-evaluable. Sperber and Wilson do not deny that there is a "relationship between truth and logic" (ibid: 72). The logical properties of conceptual representations consist in their being capable of undergoing formal logical operations --

> Conceptual representations must have logical properties: they must be capable of implying or contradicting one another, and of undergoing deductive rules. (ibid)

Furthermore, logical operations are by (Sperber and Wilson's) definition truth-preserving. But how can a "semantically incomplete logical form", like the one which (10) supposedly encodes, have logical properties -- how would we identify it as a *logical* form, or know that it logically implies or contradicts other logical forms -- if it cannot have any truth value for logical operations to preserve? How, for example, do we state the relation between (10) and (10a) -- i.e. that (10) logically implies (10a) if and only if (10a) is true

when (10) is true, and (10) is false when (10a) is false -- since both (10) and (10a) are supposed to be *incapable of being true or false*? And, most obviously, how can the non-propositional (10) contradict (10b), which *is* propositional?

Also, if (10) is an utterance whose linguistically encoded logical form is logically -- not communicatively -- developed (see 3.2) into the propositional form

(10c) Deirdre Wilson$_j$ carried the book$_i$ in her (Deirdre Wilson's$_j$) hand.

then the propositional form (10c) ought to entail the logical form of (10). But again this requires (10) to be true (and hence have a truth value) when (10c) is false.

Thus my argument is this. As long as linguistic semantic representations are regarded as *logical* forms which can undergo *logical* operations, it is impossible that they should also be incapable of being true or false. In fact there may be no reason why (10) cannot represent a state of affairs, and be true or false (depending on whether or not the state of affairs it represents exists). As Wilson and Sperber themselves point out (1981: 156), the range of the pronouns' possible referents is "explicitly given by [linguistic] semantic rules". Thus (10) ought to be perfectly true of any situation in which

(11) [*THIRD PERSON (FEMALE)*] carried [*THIRD PERSON (NEUTER)*] in the hand of [*THIRD PERSON (FEMALE)*].

-- i.e. in which *any member of the set of the first pronoun's possible referents* carried *any member of the set of the second pronoun's possible referents* in the hand of *any member of the set of the third/ first pronoun's possible referents*. (And the hand does not even have to belong to the subject of the sentence.)

It is possible that Carston might take (11) to be equivalent to

(11a) Some female entity in some domain carried at some past time something in that female entity's hand.

which she regards as the result of a "simplistic" attempt at formulating the semantic content of an obviously semantically incomplete LE:

> ... wherever you find an indexical you put in a phrase which spells out the indexical constraint in conceptual terms and wherever you spot an inarticulated constituent you use one of the family of 'some'-indefinites (*something*, *somewhere*, *sometime*, etc.) to make it visible. (1998: 69)

She raises two objections to this (ibid). Firstly, she argues that (10) and (11a) are not truth-conditionally synonymous. Indeed, (10) is not truth-conditionally synonymous with (11a), since (10) in fact entails (11a). But whether or not (10) is truth-conditionally synonymous with (11a) ought to be beside the point for Carston, since within the relevance framework it is unnecessary, because impossible, for the supposedly *non-truth-conditional* (10) to be *truth-conditionally synonymous* with (11a). Secondly, Carston claims that (11a) "cannot but be true", which she calls "a ludicrous result". It is not clear what is so ludicrous about that, especially in view of the fact that the extremely vague (11a) can always be enriched by the central inferential processes -- which are of course acknowledged by the relevance theorists to be an essential component of utterance interpretation -- to give a more specific and informative proposition. Thus I agree with Carston that

> ... the truth conditions of the vast majority of utterances depend on input from pragmatic processes... (Carston 1988: 176)

but at the same time deny that there is a non-truth-conditional linguistic semantics which is distinct from "real", truth-theoretic semantics.

In my view, the linguistically encoded logical form of (10) entails the fully propositional (11a), is entailed by the fully propositional (10c), and is itself a truth-conditional, truth-evaluable proposition. Let me just examine and compare the respective subjects of (10), (11a) and (10c). 'She' in (10) does not encode as "definite" (see Sperber and Wilson 1986a/ 95: 73) a concept as *Deirdre Wilson* in (10c), because 'she' does not in itself refer to a specific (female) individual. However, 'she' does encode a more "definite" concept that (11a)'s *some female entity in some domain*, because 'she' CAN refer to a specific (female) individual. It seems to me that, for Kempson, Carston and Sperber and Wilson, the truth-theoretic semantics of a proposition has to "represent", not only "a state of affairs in a possible or actual world" (Sperber and Wilson 1986a/ 95: 72), but a very particular and narrow state of affairs. Thus they maintain that truth-theoretic status is reached only when *she* has been inferentially enriched to the point that it refers specifically to *the* (female) individual Deirdre Wilson, or Robyn Carston, or whoever is being referred to. I, on the other hand, argue that, for 'she' to encode a semantically complete logical form, it is enough that 'she' *can* (as opposed to *does*) refer to a specific (female) individual belonging to a possible or actual world. Similarly, it is not necessary to consider any possible or actual world in order to know, for example, that (10) entails (10a) and contradicts (10b), or to confirm that a tautology such as 'War is war' or 'A bachelor is an unmarried man' is necessarily true.

For Kempson and Carston, the *distinction* which they make between linguistic and truth-theoretic ("real") semantics depends (partly) on the apparent non-truth-conditionality of linguistically encoded logical forms, while the *relation* which they assume exists between the two kinds of semantics is possible only if non-truth-conditional logical forms can undergo logical operations:

> ... the linguistic meaning of expressions has to interact with contextual parameters of a rich and complex sort to determine the truth-theoretic content of the propositions that a sentence expresses. (Kempson 1986: 77)

> It seems then that we must distinguish two kinds of semantics, linguistic and truth-conditional, the former naturally figuring only in a theory of utterance meaning, the latter taking as its domain propositional forms, whether of utterances or unspoken thoughts. Linguistic semantics IS autonomous with respect to pragmatics; it provides the input to pragmatic processes and the two together make propositional forms which are the input to a truth-conditional semantics. (Carston 1988: 176)[4]

In Carston 1998(: 64-5), the differences between linguistic and truth-theoretic semantics are set out in greater detail. Linguistic semantics has to do with the encoded meanings of natural language formulas, and is "translational" in that it involves the mapping of natural language logical forms into the formulas of the internal representational system (i.e. the language of thought). Truth-theoretic semantics is the semantics of this internal language, and is "real" rather than translational because it relates the mental representations of the language of thought to the objects and states of affairs in the world which they represent. In other words, the distinction between linguistic and truth-theoretic semantics is co-extensive with the distinction between linguistic decoding and the central inferential process.

Sperber and Wilson claim that

---

[4] One might argue that Kempson's "linguistic meaning of expressions" is exactly the same thing as her "propositions that a sentence expresses" -- sentences *are* linguistic expressions, and the propositions they express *must* constitute "linguistic meaning". Similarly, there appears to be little to choose between Carston's "utterance meaning" and "propositional forms... of utterances or unspoken thoughts". Considering that the objective of Carston 1988 is to argue for the distinction between linguistic *sentence* semantics and the truth-conditional semantics of the propositions which, as she herself points out, are expressed by *utterances*, the use of "utterance meaning" to mean 'linguistic semantics' is confusing, to say the least. Therefore one of the main problems of Kempson 1986 and Carston 1988 is the rather haphazard manner in which the terms "sentence" and "utterance" are used.

> A formula is semantically interpreted by being put into systematic correspondence with other objects: for example, with the formulas of another language, with states of the user of the language, or with possible states of the world. (1986a/ 95: 173)

According to Kempson and Carston, only truth-theoretic semantic formulae can be "put into systematic correspondence" with possible states of the world. They regard natural language formulas as non-truth-theoretic:

> Those entities which all truth-based theories of meaning assume as the basis of semantics under one conception or another -- the truth-bearing entities, propositions -- will not be generated by the grammar at all. (Kempson 1986: 101-2)

> The language ability... is the ability to map linguistic forms onto logical forms matching to a high degree the mappings made by a certain group of others (the speakers of English). In theory this ability could exist without the further capacities involved in matching these with conditions in the world. (Carston 1988: 177)

Their linguistic semantics, then, maps linguistic forms onto "*concepts*" (Carston 1988: 179, fn 16) or "*linguistic meanings*" (Kempson 1986: 100-1) -- by which they mean logical forms which cannot be true or false of states of the world or possible worlds. But, as Sperber and Wilson point out, these logical forms can be *developed* -- with "input from pragmatic processes" (Carston), or by "[interacting] with contextual parameters of a rich and complex sort" (Kempson) -- into propositional forms which CAN BE true or false of states of the world or possible worlds. It is these propositional forms which belong to the domain of truth-theoretic semantics.

The distinction between linguistic and truth-theoretic semantics would be greatly weakened should all linguistically encoded semantic representations -- or indeed any linguistically encoded semantic representation -- turn out to be perfectly propositional. As

we have seen above, Sperber and Wilson maintain that most linguistically encoded logical forms are in fact non-propositional and therefore incapable of being true or false of states of the world or possible worlds. For example, the linguistic expression 'She carried it in her hand' is claimed to be semantically incomplete, most obviously because the pronouns "do not correspond to definite concepts, but merely mark an unoccupied place where a concept might go" (1986a/95: 72-3). For Sperber and Wilson, this "non-propositionality"/ "semantic incompleteness" is one of the two features[4] of linguistically encoded logical forms which distinguish linguistic semantics from truth-theoretic semantics. I have argued that linguistically encoded logical forms must be fully propositional and truth-conditional. But even if they were not, there would still be problems with the claim that they are inferentially enriched to give the propositional forms which come under truth-theoretic semantics.

To begin with, there are two questions which have already been raised, but ought to be discussed in greater detail: what exactly does a linguistically encoded logical form consist of, and how is it (to use Sperber and Wilson's term) *developed* into a propositional form? Taking into consideration the distinction between the two kinds of semantics, the more specific and interesting versions of these questions would be: what exactly does a *non-truth-conditional, non-propositional* logical form consist of, and how is it developed into a *truth-conditional propositional form*?

According to Sperber and Wilson, a linguistic semantic representation is non-propositional because some of its constituents "do not correspond to definite concepts, but merely mark an unoccupied place where a concept might go" (1986a/ 95: 72-3). This statement seems to have some sort of internal inconsistency. The first part -- "do not correspond to definite concepts" -- appears to presuppose that the semantic constituents

---

[4] The other, to be discussed in section 3.3, has to do with Sperber and Wilson's assumption that there are fully propositional forms which nevertheless are not what "humans are interested in communicating" (Sperber and Wilson 1986a/ 95: 174).

do correspond to concepts, but only to concepts which are not *definite*. But "merely mark an unoccupied place where a concept might go" implies that the pronouns in 'She carried it in her hand', for instance, *do not correspond to concepts -- "definite" or "indefinite" -- at all*, but are merely slots in which certain concepts may be inserted.

It is possible that Sperber and Wilson actually equate "indefinite concepts" with "unoccupied places where concepts might go". Apart from the questionable coherence of such an assumption -- an "*indefinite* concept" is nevertheless a concept -- there is also the question of what makes a concept "definite" or "indefinite". Presumably Sperber and Wilson are saying that the pronoun 'she' in 'She carried it in her hand' is an "indefinite concept" at the level of linguistic semantics because it does not refer to a particular individual -- i.e. it does not correspond to a "definite concept" like *Deirdre Wilson*. But the point is that the reference to *Deirdre Wilson* occurs at the TRUTH-THEORETIC SEMANTIC level. Just because the LINGUISTIC SEMANTIC content of 'she' -- something like [*THIRD PERSON (FEMALE)*] in (11) -- has to be inferentially enriched before it can be used to refer to *Deirdre Wilson* at the truth-theoretic semantic level, does not mean that [*THIRD PERSON (FEMALE)*] is in itself "indefinite" or "semantically incomplete". My argument, therefore, is that the only sense in which [*THIRD PERSON (FEMALE)*] might be "indefinite" or "semantically incomplete" is that it falls short of what the speaker intended for it to convey. But that is not what "semantically incomplete" means, if indeed "semantically incomplete" has a meaning at all.

More recently (e.g. in Wilson and Sperber 1993), Sperber and Wilson have accepted Blakemore's (1987) distinction between CONCEPTUAL and PROCEDURAL information encoded by LEs. Whereas conceptual information consists of conceptual representations with logical and truth-conditional properties, procedural information comes in the form of procedures for constraining inferential processes. Or, as Wilson and Sperber put it (1993: 11-2), procedural information constrains the inferential construction of conceptual representations which are not, or not completely, linguistically encoded.

Pronouns, for instance, encode procedural constraints on explicatures by "guiding the search for the intended referent, which is part of the proposition expressed" (ibid: 21).

In support of this argument, Wilson and Sperber cite Kaplan's (1989) example

(12) I do not exist.

as uttered by David Kaplan. Kaplan distinguishes between the *character* and *content* of an expression: whereas the content of an expression is its "propositional constituent" (1989: 523), the character is a rule which determines the content in any given context. Thus Kaplan's character and content correspond to Blakemore's procedural constraints and conceptual information respectively. Kaplan observes that, if the 'I' in (12) is treated as encoding the concept *the speaker*, then David Kaplan's utterance of (12) would express the proposition

(12a) The speaker of (12) does not exist.

If, on the other hand, we assume that the 'I' in (12) encodes a procedural rule which states that the interpreter is to identify the referent by first identifying the speaker, then in a situation where it is uttered by David Kaplan (12) would express the proposition

(12b) David Kaplan does not exist.

So, on the assumption that pronouns encode CONCEPTUAL information:

(13)(a) The SENTENCE (12), containing the pronoun 'I' which encodes the concept *the speaker*, expresses the proposition (12a).

(b) David Kaplan's UTTERANCE of the sentence (12) expresses what the sentence encodes, (12a).

(c) Both the SENTENCE (12) and what is expressed by the UTTERANCE of the sentence (12a) are "true in any state of affairs in which [(12)] is uttered and its speaker does not exist... such a state of affairs is impossible..." (Wilson and Sperber 1993: 20).

(d) Therefore both the SENTENCE (12) and the UTTERANCE of the sentence are necessarily false.

And on the assumption that pronouns encode PROCEDURAL information:

(14)(a) The SENTENCE (12), containing the pronoun 'I' which encodes a procedural constraint rather than a concept, is semantically incomplete and does not express a proposition.

(b) David Kaplan's UTTERANCE of the sentence (12), however, does express the proposition (12b), which is inferentially derived from (12).

(c) The SENTENCE (12) is non-truth-evaluable; any UTTERANCE of the sentence (12) is false.

(d) But according to Wilson and Sperber,

> Kaplan argues that though [the procedural (12)] is false whenever it is uttered, it is not necessarily false. The proposition it expresses is true in any state of affairs in which David Kaplan does not exist. In other words, [(12)] must be understood as expressing [(12b)], not [(12a)]. (1993: 20)

I have two objections to the arguments outlined in (13) and (14). Firstly, why *must* (12) be understood as expressing (12b), and not (12a)? I see no reason why necessarily

false propositions should be undesirable, just as I see no reason why propositions which "cannot but be true" should be regarded as "ludicrous" by Carston (1998: 69).

Secondly, I disagree with Kaplan's assertion that a necessarily false proposition is expressed only in the conceptual account of (12), and not in the procedural account. Note that, whereas it is both the SENTENCE (12) and what is expressed by any UTTERANCE of the sentence (12) which are necessarily false on the conceptual account, the procedural account has it that *only* the UTTERANCE of the sentence (12) -- *not* the non-truth-evaluable SENTENCE itself -- ' . can be true or false. So if procedural (12) is *truth-evaluable only when it is uttered*, and if it is (as Wilson and Sperber themselves put it) "*false whenever it is uttered*", then -- contrary to the claims of Kaplan, and of Wilson and Sperber -- procedural (12) must be necessarily false.

Here is another way of formulating my arguments against Kaplan, Wilson and Sperber's claims about the procedural account of (12). The point which Kaplan, Wilson and Sperber are trying to make is that (12) is always false if 'I' encodes the concept *the speaker*, but can be true if the pronoun is procedurally constrained to be assigned a particular referent, e.g. David Kaplan, who does exist. But the procedural information which 'I' encodes is, as Wilson and Sperber put it (1993: 22), "an instruction to identify its referent by first identifying the speaker". This means that (12) would express (12b) *if and only if David Kaplan is the speaker*. Therefore -- since (12b) is expressed only by David Kaplan's utterance of (12), and since David Kaplan's uttering of (12) is obviously *not* a state of affairs consistent with (12b)'s description of David Kaplan's non-existence -- (12b) as expressed by an utterance of procedural (12) must be necessarily false.

It seems to me that Kaplan, as well as Wilson and Sperber, occupy a somewhat ambiguous position in respect of the relation between (12) and (12b). On the one hand Wilson and Sperber refer to (12b) as "*the proposition expressed by (12)*". On the other hand, when they claim (1993: 20) that "The proposition [(12)] expresses is true in any state of affairs in which David Kaplan does not exist", they seem to regard (12b) as an

independent proposition, completely dissociated from the fact that it is expressed by (12) if and only if (12) is perceived to be uttered, and uttered by (the existing) David Kaplan.

Returning to the relevance theorists' advocacy of the procedural treatment of pronouns: in Wilson and Sperber 1993, Kaplan's account of how the conceptual treatment of 'I' may result in necessary falsehoods like (12a) is used as an argument *for* the procedural alternative. What I have tried to show is that (12) -- as a *sentence*, or whenever it is *uttered* -- expresses necessarily false propositions on both conceptual and procedural accounts. Therefore, even on Wilson and Sperber's own terms, (12) does not constitute an argument against the conceptual treatment of 'I'.

Furthermore, if the procedural vs. conceptual debate is to be seen as turning on whether or not utterances express necessarily true or false propositions, then Wilson and Sperber should also take into consideration examples such as

(12c) I am David Kaplan.

It is on the procedural rather than the conceptual view of pronouns that an utterance of (12c) would express a proposition which is, if true, then necessarily true -- i.e. *David Kaplan$_j$ is David Kaplan$_j$* -- and, if false, then necessarily false. On the conceptual account, on the other hand, (12c) or an utterance of (12c) would express *The speaker is David Kaplan* -- a proposition which is necessarily true only if it is necessarily true that the speaker is David Kaplan.

In fact it is not clear if the procedural account of pronouns is consistent with the semantic incompleteness view, for it could be argued that Blakemore's procedures, and Kempson's grammar-specified sets of "constraints on constructing propositions" (1986: 101-2) are *themselves propositional*. Carston's remark about how one can replace indexicals with "a phrase which spells out the indexical constraint in *conceptual* terms" (1998: 69, my italics) appears to corroborate the view of Blakemore's procedural

information as sets of propositions, albeit ones which are obviously far more general than those which have been pragmatically enriched and are expressed at utterance level. Put it this way. If a proposition is thought of as defining a logical space, then a set of constraints on constructing a proposition would consist of another logical space within which the logical space defined by the proposition must be constructed. In short, the sets of constraints ought to be propositions too.

Perhaps the words which Kempson uses to define linguistically encoded logical forms -- "sets of constraints on constructing propositions" -- were chosen for the purpose of distancing logical forms from propositional forms; of stressing that logical forms, as mere *sets of constraints on proposition construction*, belong to a separate category as far as the propositions themselves are concerned. The same may be said of Blakemore's linguistically encoded *procedural constraints*. Carston, in fact, states that logical forms and propositional forms are "two different kinds of entity" (1988: 176). But she does not state *why* they are "two different kinds of entity", and it must be pointed out that the difference does not actually follow from the claim that logical forms have "unoccupied places where concepts might go" while propositional forms do not. This is because non-propositionality is conceived of as a sufficient but not necessary condition for non-truth-evaluability within this particular framework. That *most*, but not *all*, logical forms are supposed to be non-propositional --

> ... while linguistic sense makes a crucial contribution to truth conditions it *almost* never supplies a truth evaluable propositional form. (Carston, 1988: 175; my italics)
>
> ... the sense of a sentence is *often* an incomplete logical form. (Sperber and Wilson, 1986a/95: 73; my italics again)

-- seems to be one of the reasons why Carston, Kempson, Sperber and Wilson distinguish between truth-theoretic and linguistic semantics in the first place. However, it cannot

constitute a basis for the distinction itself as long as there are a number (even a very small number) of logical forms which *are* fully propositional, like Sperber and Wilson's (10b):

(10b) No one ever carried anything.

In fact, to say that *some* linguistically encoded logical forms are fully propositional amounts to conceding that *all* of them are, since they enter into logical relations with one another.

To conclude, some observations on the following remarks made by Kempson and Carston on the relation between linguistic and truth-theoretic semantics:

> Linguistic semantics is... a partial translation algorithm defined in solely syntactic terms for mapping linguistic strings which are not transparent with respect to their inferential properties on to *fully specified strings of some selected language of inference* which are. In consequence, there is no reconstruction of truth-theoretic content defined *directly* over strings of natural languages as part of the grammar... (Kempson, 1986: 102; my italics)

> Distinguishing two kinds of semantics in this way... shows further that the semantic representation of one language may be a syntactic representation in another, *though the chain must end somewhere with the formulas related to situations and states of the world or possible worlds.* (Carston, 1988: 177; my italics again)

These remarks seem to compromise the original *distinction* between linguistic and truth-theoretic semantics. The quotation from Kempson implies that, although the sets of constraints on proposition construction -- i.e. the logical forms -- specified by the grammar are not *the actual truth-evaluable propositional forms*, they still bring about the "mapping" of these logical forms onto "fully specified strings of some selected language of inference". Thus logical forms are not "*directly*" truth-evaluable but, presumably, may nonetheless be *indirectly* so in virtue of their mapping function. Meanwhile, Carston

insists that linguistic formulas must somehow be related to states of the world or possible worlds ("the chain must end somewhere with the formulas related to situations and states of the world or possible worlds"), and then promptly posits a "third kind of semantics" to relate them:

> In addition to a linguistic semantics mapping linguistic forms onto concepts and a truth-conditional semantics relating propositional forms to the real world there is a third kind of semantics, a logical or conceptual role semantics, concerned with relations of entailment, contradiction, etc., amongst logical and propositional forms. (Carston, 1988: 179, fn 16)

Also, in Carston 1998(: 64-5), it is stated that linguistic semantics "translates" natural language sentences into formulas of the language of thought, and that

> ... many natural-language expressions have a real world truth-conditional semantics *by inheritance*; that is, given that they map onto parts of propositional thought representations they can be thought of as having a truth-conditional content that those parts of the thought representations have. (my italics)

Naturally she is careful to say that LEs map only onto "parts of propositional thought representations"; that there are "a range of elements which do not encode conceptual truth-conditional content" (ibid: 65); and that the mentalese form "ijk" which the natural language form "abc" encodes is "incomplete" (ibid:64). Nevertheless, it could be said that her "third kind of semantics" and "truth-conditional semantics BY INHERITANCE" completely overturn the distinction between linguistic and truth-theoretic semantics which she had made in the first place.

3.3 Sperber and Wilson's account of verbal communication: propositional form identification

Sections 3.2 and 3.2.1 centred on one of the two reasons why a distinction between two kinds of semantics is felt to be necessary by the relevance theorists: it follows from the view of linguistic semantic representations as *non-truth-evaluable* logical forms, that linguistic semantics must have a truth-evaluable counterpart -- truth-theoretic semantics -- which consists of truth-evaluable propositional forms. In this section I discuss the other reason, which has to do with the nature of the *propositional form of an utterance* (henceforth P).

P is variously referred to by relevance theorists as *the proposition expressed by the utterance (or speaker)*, *what is said*, and, most importantly, *the truth-conditional content of the utterance*. Grice's 1967 notion of what is said was, after all, for the circumscribing of what pertains to the truth and falsity of philosophical statements. P also appears to be determined by speaker's intentions: in Carston 1998(: 26) it is defined as "that proposition which it is rational to assume the speaker intended to express".

A minimally propositional form need not necessarily be -- and in fact seldom is -- the P which Sperber and Wilson's process of logical form development produces. So the second reason for distinguishing between truth-theoretic and linguistic semantics is based on certain notions about what makes a propositional form *THE* PROPOSITIONAL FORM of an utterance. Put another way, the second reason is based on notions about the circumstances under which a propositional form belongs within the domain of truth-theoretic semantics. The central, most general, of these notions may (I argue) be summarised thus: the property which distinguishes a propositional form as a truth-theoretic semantic entity is not actually inherent in that propositional form, but depends on its capacity for serving a particular purpose within a particular context.

According to Sperber and Wilson (1986a/95: 183-93), the hearer's task is to identify the RIGHT PROPOSITIONAL FORM, i.e. the one that was intended by the speaker. Quite simply, the propositional form of an utterance is the propositional form whose retrieval requires the least effort, and which leads to an overall interpretation that is consistent with the principle of relevance. Sperber and Wilson claim that such a propositional form -- the *right* propositional form -- is itself consistent with the principle of relevance.

Propositional form identification involves three sub-tasks -- DISAMBIGUATION, REFERENCE ASSIGNMENT and GENERAL ENRICHMENT. During each of these sub-tasks, the hearer first chooses the most accessible solution, abandoning it only if it leads to a propositional form and an overall interpretation which are not consistent with the principle of relevance. Take, for example,

(15) The child left the straw in the glass.

-- which could mean that the child left the drinking tube in the glass, or that the child left the cereal stalks in the glass. Sperber and Wilson observe that

> At a purely linguistic level, there is no reason to assume that the *cereal-stalk* sense of 'straw' is less accessible than the *drinking-tube* sense; no reason, then, why one interpretation should be preferred. (ibid: 186)

And yet, in the absence of a "special context", one interpretation -- the *drinking-tube* interpretation -- is preferred. This has to do with contextual factors. As Sperber and Wilson point out:

> A child drinking from a glass with a straw is a stereotypical event which we assume... is recorded in the form of a single chunk, stored at a single location in memory and accessed in a single unit. Such a chunk constitutes a highly accessible encyclopaedic context in which the *drinking-tube* interpretation of [(15)] can be processed at minimal cost. (ibid)

Information about cereal stalks, on the other hand, is probably stored separately from information ahout children and glasses -- hence the relative inaccessibility of the encyclopaedic context required for the processing of the *cereal-stalk* interpretaion.

However, the more accessible *drinking-tube* interpretation need not necessarily be the more relevant one. Take the pronoun in Mary's utterance "It will get cold". For Peter, the most easily recoverable referent for 'it' may be the dinner -- if, for example, he knows that the dinner has been on the table for the past fifteen minutes. But let us say that Mary goes on to utter

(16) Put your jumper on.

In that context, Peter would have to abandon the *dinner* interpretation, and conclude that the pronoun is an expletive which does not refer to anything.

The inferential processes which I have discussed so far, *disambiguation* and *reference assignment*, are aimed at minimal propositionality. However, Sperber and Wilson argue that minimal propositionality is necessary but usually not sufficient for a proposition to be THE propositional form of an utterance. They claim (ibid: 189) that the interpretation recoverable from an utterance like

(17) It will take some time to repair your watch.

by decoding and reference assignment

> ... is a truism and thus irrelevant. It goes without saying that watch-repairing is a process with a temporal duration, and a speaker observing the principle of relevance must have intended to express something more than goes without saying. (ibid)

Sperber and Wilson claim that the logical form of (17) would also have to be *enriched,* through the pragmatic narrowing of the meaning of 'some time', to give a more relevant propositional form like

(18) It will take longer than expected to repair your watch.

Note that this sub-task of propositional form identification, *general enrichment,* is a crucial aspect of how Sperber and Wilson's notion of *what is said* differs from Grice's in Grice 1975. Of the utterance

(19) He is in the grip of a vice.

Grice remarks:

> One would know that [the speaker] had said, about some particular male person or animal *x*, that at the time of the utterance (whatever that was), either (1) *x* was unable to rid himself of a certain kind of bad character trait or (2) some part of *x*'s person was caught in a certain kind of tool or instrument... But for a full identification of what the speaker had said, one would need to know (a) the identity of *x*, (b) the time of utterance, and (c) the meaning, on the particular occasion of utterance, of the phrase *in the grip of a vice* [a decision between (1) and (2)]. (1975: 44)

Thus Grice's claim appears to be that what is said requires only the contributions of reference assignment ("the identity of x") and disambiguation ("a decision between (1)

and (2)"). The identification of Sperber and Wilson's propositional form, on the other hand, involves the additional sub-task of general enrichment. The relevance theorists are generally of the opinion that the results of disambiguation and reference assignment are only enough for the construction of a minimally propositional form from a linguistically encoded logical form. As Carston observes:

> ... most pragmatists working in the Gricean framework have adopted as a working principle the view that any pragmatically determined aspect of utterance interpretation, apart from disambiguation and reference assignment, is necessarily an implicature. The explanation for cutting things this way lies with the further assumption that the explicature must be truth-evaluable; so Grice and the Griceans are prepared to let in just whatever is necessary in addition to linguistically determined content to bring the representation up to a complete propositional form, i.e., something capable of bearing a truth value. (1988: 163)

Thus the disagreement between Grice and the relevance theorists over the notion of what is said consists in the following, opposing claims: that (i) what is said is linguistic content plus whatever inferentially derived meaning that is necessary for minimal propositionality (Grice), and (ii) minimal propositionality is necessary but not sufficient for something to have been *said* -- hence the need for the third sub-task, general enrichment (the relevance theorists).

As I mentioned above, Sperber and Wilson maintain (1986a/95: 189) that what is recovered from the utterance

(17) It will take some time to repair your watch.

after disambiguation and reference assignment -- the proposition

(17a) The repairing of the hearer's watch is a process with a temporal duration.

-- requires enrichment to give something like

(17b) The repairing of the hearer's watch will take longer than the hearer expected.

According to Sperber and Wilson, (17a) illustrates that "the gap between semantic representations and propositional forms cannot be closed merely by disambiguation and reference assignment" (ibid: 188-9). Carston also maintains (1988: 164-5) that the truth-evaluable proposition recovered after disambiguation and reference assignment from the similar

(20) It will take us some time to get there.

is not "THE truth-evaluable propositional form" (ibid: 177).

Sperber and Wilson's notion of the "right propositional form" and how it is identified should be considered in connection with their assumption that it is the apparent non-truth-evaluability of (most) linguistically encoded logical forms which distinguishes these logical forms from truth-theoretic semantic entities. The truth-evaluable, fully propositional form (17a) is also excluded from the truth-theoretic semantics, on the grounds that it is not the "right propositional form" of the utterance. But if (17a) is NEITHER *the "right propositional form" of the utterance* NOR, because it has already undergone disambiguation and reference assignment, *the utterance's linguistically encoded logical form*, then what is it? Carston, Sperber and Wilson do not provide an answer to this question.

Sperber and Wilson do, however, claim that (17a) "goes without saying" and is therefore "a truism" (1986a/95: 189). Carston also argues (1988: 164, my italics) that it is

"*too unspecified* to be taken as the explicature of the utterance". In short, the relevance theorists stress the fact that (17a) is NOT CONSISTENT WITH THE PRINCIPLE OF RELEVANCE. So their argument appears to be that a fully propositional form may *not* constitute an utterance's truth-theoretic semantics,[5] simply on the grounds that it is not relevant. In other words, the relevance theorists' decision about whether or not a proposition is a truth-theoretic semantic entity seems to turn largely on whether or not the proposition is (assumed to be) intended to be communicated. As Carston puts it, minimally propositional forms like (17a)

> ... are not among the constructs of [a cognitive account of communication], as they are neither communicated assumptions [i.e. explicatures] nor the vehicle by means of which these assumptions are recovered [i.e.P]. (1998: 161)

This probably explains the problematically close relation between propositional form identification and explicature recovery. That propositional form identification and explicature recovery are closely related may seem a trivial point -- from the very beginning, Sperber and Wilson make it clear that the identification of the propositional form of an utterance is absolutely necessary for the recovery of the utterance's explicature(s), and that the "right propositional form" is the most accessible propositional form which leads to an overall interpretation consistent with the principle of relevance. However, the fact that the recovery of both propositional form and explicature is governed by the principle of relevance makes Sperber and Wilson's account of utterance interpretation unavoidably circular. The propositional form of an ordinary assertion is the

---

[5] It might even be the *linguistic* semantic representation of an utterance. An example of a fully propositional form which, according to Sperber and Wilson (1986a/95: 73), constitutes the linguistic semantics of an utterance is

(10b) Nobody ever carried anything.

See section 3.2.1.

propositional form which is consistent with the principle of relevance. It is consistent with the principle of relevance because it leads to an overall interpretation which is consistent with the principle of relevance. For an ordinary assertion, the overall interpretation which is consistent with the principle of relevance actually includes an explicature which IS the propositional form of the utterance. The problem is this. In order to identify what an ordinary assertion explicates the hearer must first identify its propositional form. But, according to Sperber and Wilson, the "right propositional form" of an ordinary assertion is the proposition that is consistent with the principle of relevance -- i.e. the proposition that leads to an overall interpretation which includes the propositional form itself as an explicature. In short, the "right propositional form" simply *is* the propositional form which is explicated. This means that, rather than identifying the "right propositional form" and *then* concluding that it is explicated, the hearer identifies the "right propositional form" *on the basis of the assumption that it is what the speaker intended to explicate*.

That the relevance-constrained process of propositional form identification has to take into account the overall interpretation of the utterance -- that the identification of *what is said* is not prior to, but dependent on, assumptions about *what is communicated* -- creates even more complications in the case of non-literal utterances. This is because the "right propositional form" of a non-literal utterance is supposed to be *relevant* and yet *not communicated/ explicated.* As I pointed out in 3.2, the relevance framework has it that the propositional form P of a non-literal utterance is not itself explicated, but explicated as *constituting what is said* -- i.e. the explicature is not P but *The speaker said that P.* If Sperber and Wilson's account of propositional form identification is anything to go by, the hearer would have to take all this into account before he is able to identify the "right propositional form". This means that, whereas the "right propositional form" of an ordinary assertion is (I argue) the propositional form which the hearer assumes the speaker intended to explicate, the "right propositional form" of a non-literal utterance

would be the propositional form which is *not explicated*, but which nevertheless *constitutes what is said.*

Before further discussion and illustration of my point, here are a few more details about the role of the propositional form in Sperber and Wilson's account of non-literal utterances. According to Sperber and Wilson, an utterance is a *representation with a propositional form*, and

> ... can represent some other representation which also has a propositional form -- a thought, for instance -- in virtue of a resemblance between the two propositional forms; in this case we will say that the first representation is an *interpretation* of the second one, or that it is used *interpretively*. (1986a/95: 229)

Sperber and Wilson's principle claim is that every utterance is used as an *interpretive expression* of a thought of the speaker's, and that a hearer makes *interpretive assumptions* about the speaker's *interpretive intention* (ibid: 230-1). This is the basis of their account of *non-literalness*. An utterance is strictly literal if it has the same propositional form, with exactly the same set of logical properties, as that of the thought which it interpretively represents (ibid: 233). Hence an utterance is less than literal if its propositional form "shares some, but not all" (ibid) of its logical properties with the propositional form of the thought which it interpretively represents. That is, the utterance's propositional form RESEMBLES that of the thought: the utterance is not strictly literal because the propositional forms are not identical, but it can still represent the thought interpretively in virtue of the resemblance between its and the thought's respective propositional forms.

A thought -- especially one which is very complex -- may be best represented by such an utterance rather than by a literal but correspondingly complex utterance. Sperber and Wilson argue that

> From the standpoint of relevance theory, there is no reason to think that the optimally relevant interpretive expression of a thought is always the most literal one... The optimal interpretive expression of a thought should give the hearer information about that thought which is relevant enough to be worth processing, and should require as little processing effort as possible. (ibid: 233)

For example, if the "false but economical" (ibid) utterance

(21) I earn £800 a month.

may be used to make manifest the same assumptions as the "strictly literal and truthful" (ibid)

(22) I earn £792.32 a month.

-- then it would be (21), not (22), which is more relevant, because it is more "economical" and therefore easier to process. Of course (21) has logical and contextual implications which the speaker did not intend to convey. But, given the principle of relevance, the hearer would be expected to recognise which these are.

With reference to the examples provided by Sperber and Wilson (1986a/95: 225, 233-4), the propositional form of an utterance like

(23) John is a tiger.

would be regarded as *what is said*, the relevant propositional form, on the assumption that it has logical and contextual implications which the speaker intended to communicate. The same propositional form would be taken to be *not explicated*, on the assumption that it has logical and contextual implications which the speaker does not accept as true, and

therefore could not have intended to communicate (1986: 234) -- i.e. on the assumption that the propositional form is false.

The problem is this. According to Sperber and Wilson,

> ... all the hearer has to do is start computing, in order of accessibility, those implications which might be relevant to him, and continue to add them to the overall interpretation of the utterance until it is relevant enough to be consistent with the principle of relevance. (ibid: 234)

In other words, the logical and contextual implications which are communicated are those which involve the least effort to recover, and which are consistent with the principle of relevance. However, this is precisely how the "right propositional form" is supposed to be identified in the first place.

Take the referring expression 'John'. Reference assignment is, of course, a sub-task of propositional form identification. My point is that, on the one hand, the question of who or what 'John' refers to determines which propositional form is the "right propositional form" of (23); while, on the other hand, the sub-task of reference assignment has to take into consideration the fact that the "right propositional form" is the propositional form which gives rise to the relevant logical and contextual implications. The hearer can only take the propositional form of (23) to be explicated, and assume that (23) is an ordinary assertion, if he thinks that 'John' refers to some tiger. But if he thinks that John is in fact a tiger, then the utterance (23) would not be relevant for him, in virtue of having very few logical and contextual implications. So he cannot take that propositional form to have been explicated. Instead, he would have to assume that 'John' refers to a particular human individual, and conclude that the "false but economical" (ibid: 233)

(24) John, a particular human, is a tiger.

is the "right" but *unexplicated* propositional form of a *non-literal* utterance.

3.4 Conclusion

What I have attempted to show in this chapter is that Sperber and Wilson's distinction between linguistic and truth-theoretic semantics has nothing to do with the nature of the linguistic in itself, but is entirely based on what "natural"/ external languages are for. The fact is that relevance theory is not a *linguistic* theory but one of *human communication and cognition*, and that external languages and their linguistic semantics -- though supposedly distinct from the truth-theoretic semantics of some mental/ internal language -- are also a "tool" (Sperber and Wilson 1986a/95: 173) for the communicatory and cognitive activities of the central systems.

I have approached Sperber and Wilson's account of verbal communication in two different ways: (i) from the angle of logical form development, and Sperber and Wilson's notion of non-truth-evaluable linguistic semantic representations, and (ii) through the process of propositional form identification, and what Sperber and Wilson believe a propositional form should be. My argument is that Sperber and Wilson regard "linguistically encoded logical forms" as non-truth-conditional only because most of these logical forms have to undergo inferential enrichment before they become the sort of things which "humans are interested in *communicating*" (ibid: 174, my italics). Similarly, Sperber and Wilson's truth-theoretic *propositional forms of utterances* must be more than minimally propositional -- they have to be consistent with the principle of relevance. In my view, it does not follow from the fact that (11) or (17a) --

(11) [*THIRD PERSON (FEMALE)*] carried [*THIRD PERSON (NEUTER)*] in the hand of [*THIRD PERSON (FEMALE)*].

(17a) The repairing of the hearer's watch is a process with a temporal duration.

are seldom assumed to be communicated or worth communicating, that their semantics are thereby not "real".

In assuming that the linguistically encoded logical forms of utterances are non-truth-evaluable, and that propositional forms of utterances must be consistent with the principle of relevance, Sperber and Wilson in fact create problems for their own account of verbal communication. On the one hand, the need to construct propositional forms from linguistically encoded logical forms raises the question of how truth-evaluable logical forms can be "developed" from non-truth-evaluable ones -- not to mention the question of whether there really are such things as non-truth-evaluable logical forms. On the other hand, the stipulation that propositional forms must be consistent with the principle of relevance has the effect of collapsing the propositional form with the similarly relevance-governed explicature.

I also argue that Sperber and Wilson's associating of propositional forms with the property of relevance creates problems for their account of non-literalness. In the first place, Sperber and Wilson's definition of a non-literal utterance as an utterance whose propositional form constitutes what is said but not what is explicated adds a new and problematic aspect to the relation between propositional form and explicature. In the case of ordinary assertions, the propositional form of an utterance is identified on the grounds that it is the proposition which the hearer assumes that the speaker intended to communicate. But it is more difficult to explain why the hearer identifies a particular proposition as the "right propositional form" of a non-literal utterance, since that proposition cannot be assumed to be communicated, and therefore cannot actually be described as relevant. In other words, the problem is this: how does the hearer conclude from the assumption that the proposition is not intended to be explicated, that it is in fact

the "right propositional form" of the non-literal utterance -- the propositional form which leads to an overall interpretation that is consistent with the principle of relevance?

Actually it is no accident that the concept of the propositional form is such a problematic one, since it appears to be something whose existence the relevance theorists postulated for certain communicative and cognitive purposes. I have argued that any linguistically encoded logical form is in fact semantically complete in itself, and does not have to be *developed into a propositional form* to be truth-evaluable. In that respect, at least, propositional forms seem superfluous.

4

The "misuse" of language

4 Introduction: phonetic vehicles and "linguistic" conventions

In section 2.3 I discussed Chomsky's (1992a, 1995a) comments on what is generally regarded as the *misuse* of language. For Chomsky, this is the misuse of *I-language*, which he distinguishes from the use or misuse of non-(I-)linguistic expressions like the constructions of naturalistic inquiry (e.g. 'elm' in the botanist's sense, or 'mass' in the physicist's sense) and of social practice (e.g. the community norms of 'disinterested' and 'uninterested' meaning *disinterested* and *uninterested* respectively). This chapter also addresses the question of whether what is (mis)used is linguistic, but in relation to the more general problems which the phenomenon of *language misuse* creates for the corresponding notions of *language use* and *linguistic conventions*.

In the last two chapters I have been discussing the idea that linguistic expressions and their semantics are independent of the sort of truth-theoretic propositional meaning which humans entertain and communicate -- i.e. the sort of truth-theoretic propositional meaning which may be regarded as literal (or non-literal). For Chomsky, as well as for the relevance theorists, this truth-theoretic propositional meaning is not expressed by linguistic expressions in themselves, but by the USE of linguistic expressions. I have also argued that Chomsky and the relevance theorists fail to explain how exactly linguistic expressions are "used" -- how they constrain, or are involved with, but remain absolutely distinct from, the truth-theoretic propositional meaning which they are used to express.

The notion of *THE correct use* versus the *misuse* of a particular linguistic expression presupposes that the linguistic semantics of, say, the linguistic expression 'flaunt' has a truth-theoretic counterpart *flaunt* which constitutes the literal meaning that is expressed

when the linguistic expression is "correctly" used. In the case of the misuse of 'flaunt', by contrast, it is generally assumed that the linguistic expression 'flaunt' *is* used, but not to express the truth-theoretic literal meaning *flaunt* -- and not in accordance with the linguistic semantics which is supposed to determine what the literal meaning is. For example, someone might utter 'flaunt' to express *flout*.

But while the linguistic semantics of 'flaunt' may be seen as determining the truth-theoretic literal meaning *flaunt*, it is only the truth-theoretic *flaunt* which the language use or misuse has conscious access to. Significantly, linguistic meaning is referred to as literal meaning (and vice versa) at some point in every one of the three accounts of misuse which I cite in this chapter -- Davidson 1981 and 1986; Dummett 1986; and even Carston 1998, a large part of which is devoted to the distinction between linguistic and truth-theoretic semantics. It is usually assumed that the language user knows that *flaunt* is the literal meaning of 'flaunt', not because he knows that *flaunt* corresponds to the "correct" linguistic semantics, but because he knows some convention which states that the PHONETIC FORM of 'flaunt' is the "correct" vehicle of the literal meaning *flaunt*.

In Carston 1998 (discussed in 4.1), it appears to be the phonetic (or at least E-instantiable) form produced by the language user which constitutes necessary and sufficient evidence for a particular linguistic expression to have been used. Furthermore, Carston takes it for granted that there are definitive sets of linguistic conventions which specify that particular phonetic forms belong to (and are thereby indicative of) particular linguistic expressions and particular (truth-theoretic) literal meanings, regardless of what meaning the individual language user may have intended a certain phonetic vehicle to carry, or believed was carried by a certain phonetic vehicle. Thus, for Carston (as well as for Dummett), the speaker who UTTERS the phonetic form of 'flaunt' with the intention of expressing *flout* has misused the linguistic expression that is individuated by the phonetic form produced, in virtue of uttering/ using it to express a concept which, according to some linguistic convention, is not the meaning carried by that phonetic form.

I, on the other hand, argue that the use of a vehicle-meaning pair -- vehicle-meaning pairs being what linguistic expressions are generally considered to be (e.g. by Chomsky, the relevance theorists, Dummett) -- must involve both the production of the vehicle AND the intention to express the meaning which the vehicle is associated with. If an individual did not intend to express that meaning, then it seems to me that he has not USED that particular vehicle-meaning pair, even if he did produce the vehicle.

Thus I am much more in sympathy with Davidson's 1986 account of malapropisms (set out in section 4.2), in which he assumes that an expression is individuated by its meaning rather than its vehicle. Like Dummett (discussed in 4.3), Davidson believes that the conventions which govern vehicle-meaning relations are SOCIAL conventions. But whereas Dummett equates the linguistic with the social, Davidson argues that such social conventions are NOT (in themselves) LINGUISTIC, and that they may be deviated from, created or replaced in any context. On this view, the person who utters 'flaunt' with the intention of expressing *flout* has not so much deviated from a convention which states that 'flaunt' means *flaunt*, as adhered to a convention which states that 'flaunt' means *flout*. It follows from this that he has not misused an expression 'flaunt' with the meaning *flaunt*. Instead, he has "correctly" used 'flaunt', which, according to the convention which he intended to follow, means *flout*.

In chapters 5 and 6 I shall give a *non-linguistic* account of literal meaning and "language use/ misuse". More importantly, perhaps, I shall try to explain how Davidson's non-linguistic social conventions are related to the linguistic. This chapter, however, is mainly concerned with the difficulties of reconciling so-called "linguistic" conventions with the beliefs and intentions of language users who apparently "misuse" linguistic expressions.

4.1 Carston's ad-hoc concepts: a pragmatic account of misuse and slips

Carston 1998 (: 17-8) contains a brief discussion of what she calls "the phenomenon of misuses and slips of the tongue". Usually the speaker intends the hearer to access *what her utterance linguistically encodes*, and also *what is said* (or the *propositional form* of the utterance). This is supposed to apply even to cases like metaphor, where the speaker may not intend to *communicate* the linguistically encoded concepts but nevertheless intends her hearer to access them "in the process of arriving at the intended interpretation" (ibid: 17). In the case of a misuse or slip, however, what seems to happen is that "some of the meaning encoded in the utterance falls outside any intention the speaker has in producing the utterance" (ibid). So, with regard to her example of a speaker uttering

(1) Mary is one of the hoi polloi.

to communicate that Mary is a member of the upper classes, Carston claims that there is a disjunction between what is (linguistically encoded and) said --

> Arguably, the proposition expressed by this utterance, what is (strictly and literally) said, is that Mary is one of the common people... (ibid:18)

-- and anything which the speaker may intend the hearer to access. And this, Carston suggests, is because the encoding of the concept *common people* by the expression 'hoi polloi' is "*not a part of her [the speaker's] system of linguistic knowledge*" (ibid, my italics); and, not being what she knows, cannot therefore be something she intends the hearer to access.

From this account of (1) as an instance of a misuse or slip it is possible to infer certain points about how Carston conceives of what is linguistically encoded, and what is said. In

respect of *what is linguistically encoded*, Carston's account presupposes a "system of linguistic knowledge" that is quite independent of the speaker. This system of linguistic knowledge specifies that 'hoi polloi' linguistically encodes the concept *common people*, regardless of (i) the speaker's intention that her hearer should recover, through linguistic decoding, the concept *upper classes*, (ii) her belief that her utterance of 'hoi polloi' encodes the concept *upper classes*, and, indeed, (iii) the entire system of linguistic knowledge which she, the individual speaker, possesses. Furthermore, if it is not just what is linguistically encoded which falls outside the speaker's intentions, but also *what is said* (i.e. the propositional form of the utterance), then it would appear that in recovering the latter the hearer must sometimes go against his beliefs about the speaker's intentions. So the speaker may believe that 'hoi polloi' encodes *upper classes*, and the hearer may know this, and believe that the speaker intended to communicate that Mary belongs to the upper classes. Nevertheless, Carston argues that the hearer would come to the conclusion that "what is (strictly and literally) said, is that Mary is one of the common people". That the hearer's recovery of what is said should "fall outside the speaker's intentions", and even outside the hearer's beliefs about the speaker's intentions, seems to contradict the relevance-theoretic account of propositional form identification (see section 3.3). As described in Sperber and Wilson 1986a/ 95, all the procedures of propositional form identification -- reference assignment, disambiguation, etc -- are based mainly on the assumption that the speaker is observing the principle of relevance, i.e. that the speaker has certain *intentions*, the central of which is to aim for optimal relevance.

Carston concedes that examples like (1) seem "special and marginal", and

> ... bring an unclarity into the concept of what is said, since up to now we have been assuming that what a speaker says by an utterance is not at odds with what she believes her words to mean... (ibid)

Later, however, she reconsiders such examples in the light of her proposals for revising the relevance account of pragmatic processing. To begin with, she suggests that the process of LOOSENING should be brought in line with ENRICHMENT (ibid: 316-35, Carston 1996). According to Sperber and Wilson 1986a/95 (see also section 3.3), enrichment is one of the ways of arriving at what is said (or the proposition expressed). Loosening, on the other hand, is seen as contributing only to implicatures. Briefly, the process of loosening goes like this. The hearer decodes the loosely-used lexical concept, thereby gaining access to the logical and encyclopaedic entries for that concept. He then sorts through the sets of assumptions which constitute these entries, rejecting the irrelevant assumptions (including the ones which contradict one another), and retaining as implicatures those which are relevant in the given context.

Carston introduces the notion of AD HOC CONCEPT (1998: 310). An ad hoc concept is a relevant, communicated concept arising from the pragmatic enriching or loosening of a linguistically encoded lexical concept. For Sperber and Wilson (1986a/ 95), only enriched ad hoc concepts are incorporated into what is said, while loosened ad hoc concepts contribute only to what is implicated. Carston, however, suggests three ways of "symmetrifying" enrichment and loosening. The first symmetry position (1998: 314-6), which she rejects, has to do with bringing enrichment in line with loosening by assuming that the results of enrichment emerge as implicatures rather than contribute to the proposition expressed. The second symmetry position (ibid: 316-20) brings loosening into line with enrichment, with ad hoc concepts arising from loosening also incorporated into the proposition expressed. For example (ibid: 316), the propositional form for

(2) This steak is raw.

would be

(3) [This steak]$_y$ is raw*.

where *raw** represents a loosening of the lexical concept *raw*.

Before I go into the third symmetry position, here are four possible objections to the second symmetry position which Carston considers and rejects (ibid: 316-20). On the assumption that 'bald', for instance, encodes *totally hairless*, objection [1] is that the lexical concept would almost always be loosened, and therefore would almost never be literally used. The same goes for 'hexagonal', which, according to Carston, describes "a property of a perfect abstract form not actually found in nature" (ibid: 317). Carston, however, points out that this is actually a useful feature of language. Given that the cognitive capacities of humans -- including

> ... the ability to attribute complex mental states (such as higher order beliefs and intentions) to conspecifics, the ability to draw inferences from newly impinging stimuli by placing them in a context of existing assumptions, and the ability to recognise conceptual and other resemblances from a range of points of view... (ibid)

-- can be relied upon to carry out the appropriate pragmatic processes, it is enough, and in fact more cognitively efficient, for many of the encoded concepts to be "quite minimal" (ibid).

Under the second symmetry position, there is no question of whether the propositional form of an utterance is communicated, or embedded in an assumption schema to give a higher-level explicature. Since it would incorporate any ad hoc concept constructed in the course of interpreting the utterance, and since an ad hoc concept is always communicated, the propositional form itself must always be communicated too. In short, the propositional form must be a base-level explicature. Objection [2] is that the incorporation into the utterance's propositional form of ad hoc concepts previously assumed to contribute only to implicatures would remove the distinction between the

propositional form and that of the thought which the speaker intended his utterance to convey -- a distinction which is crucial to the relevance-theoretic account of non-literal utterances, and bridged by the relation of interpretive resemblance. This leads us to objection [3], which is that the speaker whose thought and utterance always have the same propositional form must be observing some Gricean maxim of literalness or truthfulness.[1]

As Carston herself remarks, it does follow from the second symmetry position that "the speaker always endorses the proposition her utterance expresses" (ibid: 318) -- i.e. that the proposition expressed is always (explicitly) communicated. However, she claims that there remains a relation of interpretive resemblance between *concepts in the thought of the speaker* and *communicated concepts* (including, or possibly constituted by, the proposition expressed). Note that she distinguishes quite clearly between

(A) what the speaker *intends* to communicate (the speaker's thought);

(B) what the speaker *actually manages* to communicate (what explicatures and implicatures the *hearer* recovers); and

(C) the proposition expressed by the speaker's utterance.

The second symmetry position may align (C) with (B), but does not change the relation between (B) and (A). It would appear that neither (B) nor (C) ever replicates (A), whether one takes the standard or "symmetrical" line on utterance interpretation. As I mentioned in the previous chapter, communication within the relevance framework does not result in the duplication of thoughts. Many aspects of thought are not linguistically encoded -- hence Carston's support for the "strong essentialist view" of underdeterminacy:

---

[1] Or rather, if the hearer always interprets an utterance as having the same propositional form as the speaker's thought, then it must be on the basis of an assumption that the speaker is observing some maxim of literalness or truthfulness.

> ... underdeterminacy is an essential property of the language/thought relation and no sentence ever fully encodes the thought it is used to express. (ibid: 27)

It is up to the obviously non-telepathic hearer to pragmatically infer the unencoded elements of what the speaker intended to communicate. Of these unencoded elements, ad hoc concepts seem the most reliant on the hearer's inferential efforts, and therefore the most indeterminate. What contextual information is chosen to enrich a lexical concept, which subset of a lexical concept's extension is the narrowed version, what assumptions are retained from a lexical concept's encyclopaedic entry as the loosened concept -- all this is constrained by the universal search for relevance, but also depends on the individual hearer, his encyclopaedic entry for the lexical concept, and the context of utterance and interpretation. Thus, in Carston's words, the speaker's thought is "seldom, if ever, perfectly replicated" (ibid: 18) in the hearer's mind.

Returning to objection [2], Carston's reply is this: loosened concepts incorporated into (C) lead to (C) being explicitly communicated, but -- being themselves indeterminate -- do not bring (C) any closer to (A). In fact one might say that the indeterminacy of the loosened concepts, previously the property of implicatures, is passed on to the (C) and to the base-level explicature which (C) becomes.

With regard to objection [3], Carston points out that the Gricean maxim of truthfulness concerns the relation between *what is linguistically encoded* and what is communicated: on the assumption that the speaker is observing the maxim, the hearer equates what the speaker's utterance linguistically encodes with what the speaker intended to communicate, thereby giving the utterance a *literal* interpretation. This obviously does not apply to the second symmetry position, which has to do with the relation between *what is said, or the proposition expressed*, and what is communicated. In any case, the addition of loosened concepts to the proposition expressed means that the latter "[departs]

even more radically than before [when only the enriched concepts were included] from the literal linguistic content in the logical form of the utterance" (ibid: 18).

Objection [4] (ibid: 319) concerns the notion of "the proposition literally and strictly expressed", "what is said", or "the truth-conditional content of the utterance". If the proposition expressed by the utterance of

(4) Bill is a bulldozer.

is

(5) Bill_x is a bulldozer*.

(where *bulldozer** is the ad hoc, loosened version of the lexical concept *bulldozer*), then the utterance of (4) would be true (if and only if Bill has the properties which correspond to the loosened concept *bulldozer**). Objection [4] is that (4) being true goes against our intuitions that the proposition expressed by the utterance is the obviously false *Bill is a bulldozer*, and that

(6) Bill is not a bulldozer.

is true. Carston argues that such intuitions are not altogether reliable. She observes that it is possible to "agree or disagree with someone who utters [(4)]... or even say 'that's **true**' or 'that's **not true**'" (ibid). She also points out that some hearers would say that (4) is true, and some, that (4) is false, even if all of them agree that Bill is an aggressive and domineering person. Carston maintains that these different responses correspond to intuitions coming from distinct sources:

> ... the 'true'-sayers would be considering the proposition they take the speaker to be expressing and endorsing (i.e. an explicature), while the 'false'-sayers would be tapping their knowledge of linguistic meaning, the literal encoded conceptual content of the utterance. (ibid: 331)

-- i.e. what is true for the 'true'-sayers is the propositional form which contains the loosened concept and is explicitly communicated, while what is false for the 'false'-sayers is a propositional form with the linguistically encoded lexical concept *bulldozer*. Similarly, (6) could be regarded as true for two different reasons. The hearer may assume that it is the lexical concept *bulldozer* which is being negated, and agree that it is true that Bill is not an inanimate object. On the other hand he may see (6) as a denial that the loosened concept *bulldozer** applies to Bill:

> (7) Bill is not really a bulldozer; in fact he's quite insecure. (ibid: 319)

I come now to Carston's third symmetry position (ibid: 328-30), which states that the results of enrichment and/ or loosening must be built into the propositional forms of only SOME utterances, in order for the hearer to recover (what he assumes to be) the intended interpretation. Thus the third symmetry position has it that there are other utterances which do not require the incorporation of enriched and/ or loosened ad hoc concepts into their propositional forms. According to Carston, (4) belongs to this second category of utterances. Carston cites Papafragou 1995, in which the symmetrical treatment of utterance interpretation is applied to examples of referential metaphor like

> (8) The wilting violet has finally left.

but not to cases of metaphorical predication like

(9) Maria is a nightingale.

and, of course, (4). With regard to (8), Carston agrees with Papafragou that the construction of an ad hoc concept from the encoded *wilting violet* concept is necessary for reference assignment to proceed. It would be from the extension of the loosened concept ("of a certain sort of shy, retiring person, let us suppose" (Carston 1998: 327)) that the hearer picks out the individuating concept of the person referred to. But Carston and Papafragou claim that the classic relevance theory account of non-literalness (Sperber and Wilson 1986a/ 95; see also my section 3.3) not only works for (9), but is more cost-effective: instead of first having to build an ad hoc concept into the propositional form, the hearer can pick out properties intended to be ascribed to Maria directly from the encyclopaedic entry of the lexical concept *nightingale*.

Carston argues that the third symmetry position -- that ad hoc concepts should be incorporated into propositional forms only when necessary -- is consistent with the principle of relevance, according to which the maximum number of contextual effects must be balanced with the minimum amount of processing effort. She also points out that in certain cases it may actually be impossible to build ad hoc concepts into the proposition expressed: in sustained metaphors, for example --

(10) Life's but a walking shadow, a poor player that struts and frets his hour upon the stage, and then is heard no more... (Shakespeare, *Macbeth* V.v.24-6)

-- and the metaphorical use of entire sentences --

(11)(a) The cracks are beginning to show.

(b) The lion is roaring again. (These examples and (10) from ibid: 328)

But although Carston states (ibid: 328, 329) that the classic relevance theory account of non-literalness is satisfactory for such examples, it is quite clear that the "third symmetry position" treatment she suggests for, say,

(4) Bill is a bulldozer.

deviates from Sperber and Wilson 1986a/95 in more ways than one. Where ad hoc concepts are not built into the proposition expressed, as in the case of (4), Carston's account has instead a propositional schema like

(12) $Bill_X$ is [ ]

with reference assigned and the lexical concept in the predicate "functioning as a constraint on the sorts of concepts that may occupy the open slot" (Carston 1998: 333). It follows from this that "there simply is no complete proposition expressed by the speaker or constructed by the hearer" (ibid: 333) -- in fact, that "a fully propositional form at the *explicit level* need never be entertained" (ibid: 329, my italics). Thus the third symmetry position reverts to the classic relevance account of the loosening process, but adheres to the "symmetrical" conflation of propositional form and explicature[2] -- even where the propositional form to be explicated is *not fully propositional*.[3]

---

[2] It has to be mentioned that this notion -- of the proposition expressed always being explicated -- comes only with the second and third symmetry positions, and is completely inconsistent with the first.

[3] Note that Carston's "third symmetry position" account of *enrichment* is even more radical. The idea that no (fully) propositional form is explicitly communicated is not that remote from the classic account of loose use in metaphor, according to which the (fully) propositional form is not itself communicated but has to be embedded in a higher-level explicature. But it certainly is a departure from the standard notion of enrichment: while Sperber and Wilson view enrichment as a way of conceptually adding to the encoded logical form, Carston suggests that it may result in a less propositional, in fact non-propositional, "propositional schema" like (12) in the main text.

Here I list what seem to me to be the main innovations of Carston's third symmetry position:

(A) The results of both loosening and enrichment may or may not be built into the proposition expressed (depending on whether ad hoc concepts are necessary for arriving at the intended interpretation of the utterance).

(B) If they are, then obviously the proposition expressed would contain ad hoc concepts arising from loosening or enrichment.

(C) If they are not, the proposition expressed -- or rather, the "proposition" expressed -- would turn out to be a non-propositional "propositional schema".

(D) Whichever the hearer recovers (fully propositional forms with ad hoc concepts, or non-propositional "propositional schemas"), it is assumed to be explicitly communicated.[4] In other words, the third symmetry position dispenses with the classic relevance theory notion of proposition expressed or "what is said", with the result that "the only level in the whole process of utterance interpretation that is not communicated is the logical form" (ibid: 319).

Carston's third symmetry position may be summed up thus: to *either* of the levels of utterance interpretation which are communicated (i.e. to either explicature or implicatures) may be contributed ad hoc concepts resulting from *either* enrichment or loosening -- depending, of course, on which procedure is most consistent with the principle of relevance in a given context. Within the third symmetry framework, it is equally possible for the main import of an utterance to be implicitly communicated as it is for it to be

---

[4] Actually there appears to be no reason for positing "propositional schemas" in explicatures, except for the possibility that -- if it contained *the lexical concept itself* -- the proposition expressed could NOT be what the speaker intended to communicate. That, of course, is what happens in the standard account of non-literalness: for instance, the propositional form of (4) is not explicated because Bill is a human being, and not compatible with what 'bulldozer' linguistically encodes.

explicitly communicated. For (8) it would be carried by the explicature, a proposition containing ad hoc concepts. For (9) and (4) it would lie in the implicatures, for which the explicature -- a non-propositional "propositional schema" -- would function merely as a "vehicle".

It should be clear by now that one of the most important implications of the third symmetry position has to do with the distinction between explicatures and implicatures, the two types of communicated assumptions. Carston argues that it is a mainly *derivational* distinction -- explicatures are derived by developing the linguistically encoded logical form, and implicatures by the inferring of whole new assumptions (ibid: 332, 335) -- and has little to do with notions of "what is strictly and literally said". This derivational difference is further reduced within the third symmetry framework, according to which both logical form development and purely pragmatic inference involve the construction of ad hoc concepts through loosening or enrichment. Thus Carston claims that the distinction between explicatures and implicatures is NOT one

> ... which is of great importance for a hearer, at least if it is viewed as a representational distinction which is supposed to have some impact on the way he views the set of communicated assumptions. (ibid: 332)

How, then, does the third symmetry position account for the phenomenon of misuses and slips? Carston's only comment regarding this issue is that slips of the tongue like

(13) She always flaunts the rules.

> ... will be correctly interpreted, that is, interpreted in line with the speaker's informative intention, although she hasn't produced the best

possible linguistic evidence to ensure fulfilment of that intention. (ibid: 334)

In the rest of this section I shall attempt to expand on this single statement.

It seems to me that (D) -- Carston's assertion that there is no level of proposition expressed distinct from that of explicature, and that "a fully propositional form at the explicit level need never be entertained" (ibid: 329) -- is central to the apparent resolution of cases of misuse and slips within the third symmetry framework. One of the features of misuse/ slips which makes them seem "marginal" is that what is said in those cases appears to fall outside the speaker's intentions. For the third symmetry position, there is no question of the proposition expressed falling outside the speaker's intentions, since the proposition expressed is always the one which the speaker intends to communicate.

The question, though, is: what does the speaker intend to communicate, and how is it recovered by the hearer? Firstly, it is clear that in the case of

(13) She always flaunts the rules.

what is intended to be communicated, and "correctly interpreted", is something like

(13') She always flouts the rules.

But does the hearer arrive at (13') by incorporating ad hoc concepts into the proposition expressed, or through a propositional schema? In other words, is the main import of an utterances like (13) explicitly or implicitly communicated? This question may seem unnecessary in the light of the third symmetry position, according to which strength of communication is more fundamental than the means. However, one could argue that the hearer is not likely to recover (13') in the form of implicatures arising from

(14) She always [ ] the rules.

-- the reason being that it seems impossible for what is communicated to be "read off the lexical concept" (as Carston puts it, ibid: 328) retained in (14) as a "constraint" on what the open slot may contain. While the hearer of

(9) Maria is a nightingale.

is able to pick out, say, *a wonderful singer* from the encyclopaedic entry of the lexical concept *nightingale*, there appears to be no way the concept of *flaunt* can lead one to *flout*, or *hoi polloi* to *the upper classes*.

At this point it may be useful to compare misuses and slips with all the other examples which have been discussed so far, e.g.

(8) The wilting violet has finally left.

(9) Maria is a nightingale.

For the latter -- even for the most non-literal of them -- the recovery of what is communicated involves the operation of processes (enrichment, loosening) upon lexical CONCEPTS. For an utterance like (13), on the other hand, one might say that no amount of conceptual enriching or loosening could induce the lexical concept *flaunt*, with its logical and encyclopaedic entries, to point the hearer in the direction of the concept which the speaker intended to communicate. This may be what Carston means when she says that the speaker who has a slip of the tongue "hasn't produced the best possible linguistic

evidence" (ibid: 334) for her communicative intentions.[5] For instead of leading the hearer along the relatively well-defined and well-established relations between CONCEPTS, as it usually does, the linguistic evidence in cases like (13) provides access to the speaker's intentions through the rather more hit-and-miss notion of PHONETIC -- i.e. physical -- RESEMBLANCES. In short, the lexical *concept* cannot and does not constitute the linguistic evidence for the inferring of what is communicated -- it is the *phonetic properties* which have that role.

This leaves us with the possibility that (13') is explicitly communicated, and also that the procedure of arriving at the explicature via the construction of ad hoc concepts cannot involve the loosening or enrichment of lexical concepts. Instead, the hearer may have to pick a phonetic form similar to the one in the linguistic expression uttered -- [flaʊt], say, for its resemblance to [flɔ:nt], and its encoded concept's compatibility with the direct object *the rules* -- and then incorporate the new lexical concept as an ad hoc concept in the proposition expressed. Furthermore, it is not even clear if the explicature should be represented as

(15) She always flaunts* the rules.

for the reason that the relevant ad hoc concept was not, could not have been, directly derived from the lexical concept *flaunt*. Another argument against the 'flaunt*' in (15) is that the ad hoc concept constructed by the hearer of (13) seems far more determinate, more definite, than those in the metaphor examples, like *wilting violet* in (8) -- the ad hoc concept for (13) appears, quite unequivocally, to be the concept encoded by 'flout', i.e. *flout*. So, if the lexical concept *flaunt* could be said to undergo any process, it would be

---

[5] As I mentioned in section 3.1.1, the sole purpose of language within the relevance framework is to provide *evidence* from which hearers infer the intentions of speakers (Sperber and Wilson 1986a/ 95: 27, 163).

the process of loosening taken to the limit -- the dropping of, not some, but *all* of the elements of the linguistically encoded content.

The same could be said of the uttering of

(1) Mary is one of the hoi polloi.

to communicate

(1') Mary is one of the upper classes.

The fact is that the concept ~~*upper classes*~~ common people is not something which the hearer was supposed to retrieve in the first place. As Carston points out, the encoding of *common people* by 'hoi polloi' is "*not a part of [the speaker's] system of linguistic knowledge*" (ibid:18, my italics); and, not being what he knows, cannot therefore be something he intends the hearer to decode. Moreover, the phonetic properties of 'hoi polloi' may also play a part in the interpretation process -- Carston suggests that there could be a "sound association" with 'hoity toity' (ibid: 17)

Thus I argue that Carston's third symmetry position does not solve the problem presented by the phenomenon of misuse and slips -- the problem of *what is said* falling outside the speaker's intentions -- simply because it does not appear to offer a solution for the more fundamental problem of *what is linguistically encoded* falling outside the speaker's intentions, and in many cases outside the speaker's beliefs and his "system of linguistic knowledge". As Carston herself puts it,

> ... up to now we have been assuming that what a speaker says by an utterance is not at odds with *what she believes her words to mean*... (ibid: 18, my italics)

Obviously there are cases of misuse where the speaker does know what the utterance encodes, i.e. misuses which are slips. The person who utters (13), for example, may know that 'flaunt' encodes *flaunt* and not *flout*, but utters 'flaunt' due to what Carston calls a

> ... temporary [malfunction] of the system, brought on by performance factors such as tiredness or emotional strain. (ibid)

The important point, however, is that all the examples involve the encoding of lexical concepts which the speaker did not intend the hearer to access/ decode, did not intend to use, and in fact did not even have in mind -- either because the encoding of the concept by the expression uttered is "not a part of the speaker's system of linguistic knowledge", or because the speaker knows that the expression encodes the concept but has no intention, and merely commits the "articulatory error" (as Carston calls it, (ibid)), of uttering the expression. Henceforth I shall refer to the first sort of speaker as speaker M (for "misuse pure and simple"), and the second as speaker S (for "slip").

This brings me to the main argument of this section: if the speaker did not intend the concept encoded by some linguistic expression to be accessed -- by *anybody*, including the speaker himself -- then surely he cannot be said to have USED that concept, the encoding of that concept, and hence *the particular linguistic expression constituted by the encoding of that concept*. So I argue that speakers M and S, in failing to entertain the semantic properties of the linguistic expressions which they are supposed to have used, do not actually use those linguistic expressions at all. And if they do not use the linguistic expressions, then they do not MISUSE them either.

Carston's account of misuse/ slips, on the other hand, presupposes that to UTTER a linguistic expression IS to use it. In ordinary cases of language use (including metaphor, irony, etc), as well as in cases of misuse involving speakers M,

(A)(i) what the speaker uses, and

(A)(ii) what the speaker utters

are generally regarded as the same. For such cases, Carston assumes that (A)(i), together with

(B) what the interpreter decodes, and

(C) what the speaker intends to use,

refer to the same linguistic expression. For slips, and for speakers S, Carston would presumably take (A)(i) and (B) to refer to the same linguistic expression (e.g. 'flaunt'), while the "slipping" speaker would have originally intended to use another linguistic expression (e.g. 'flout').

Carston's equating of (A)(i) with (B) (and, in the case of speakers/ hearers M, with (C) as well) hinges upon her assumption that (A)(i) *IS* (A)(ii). Quite simply, if what the speaker uses is what he utters, and what he utters is the physical entity which the interpreter perceives and decodes, then (A)(i) must be (B). More importantly, misuse cases are almost always discussed in terms of how they are "correctly" interpreted, with the hearer tacitly set up as the ideal language user with some definitive set of linguistic rules on his side. And it is these linguistic rules, of course, which are the crucial factor in the identifying of the linguistic expression used, and even intended to be used, by the speaker. Note that, with the shift of the proposition expressed to the level of explicature under the third symmetry position, literalness for Carston is no longer a property of what is said (as it was for the classic relevance account of utterance interpretation), but is located in "*the literal encoded conceptual content of the utterance*" (ibid: 331, my italics). That there is such an entity as THE literal linguistically encoded content, as determined by the rules of the language, is obviously presupposed by the notion of misuse/ slips. Without

the notion of standard literal meanings which are independent of the individual speaker/ hearer, no expression could ever be said to be misused. Clearly, for Carston, the question of what linguistic expression is used and/ or misused is not decided by the speaker, even though it is the speaker who uses it, but by the linguistic rules which connect the physical entity produced by the speaker to its literal linguistically encoded meaning.

Of course the speaker who utters, say, (13) might happen to agree with that set of linguistic rules, and with the hearer, about what linguistic expression he has used. Speaker S, who believes that 'flaunt' encodes *flaunt* and whose utterance of (13) was an "articulatory error", possibly believes that he has unintentionally used -- and, in doing so, misused (or rather, mistakenly used) -- 'flaunt' when he actually intended to use 'flout'. As for speaker M, who believes that 'flaunt' encodes *flout*: it may be that Carston's claim about the misuse cases, that they involve what is linguistically encoded falling outside the speaker's intentions, implies that *what remains of the linguistic expression within the bounds of the speaker's intentions* -- most conspicuously, the phonetic properties -- is sufficient for the linguistic expression (i.e. 'flaunt' encoding *flaunt*) to be described as being used by the speaker. In contrast, it is not just what is linguistically encoded but every other element of the linguistic expression which falls outside speaker S's intentions -- he did not intend to utter that linguistic expression, and he is aware that he has uttered what he did not intend to utter.

Thus both speakers S and M are regarded by Carston as using the encoding of *flaunt* by 'flaunt' -- M, solely in virtue of *intending* to produce, and producing, the phonetic form [flɔ:nt]; S, solely in virtue of *unintentionally* producing the phonetic form [flɔ:nt] (which he happens to believe encodes *flaunt*).The concept *flaunt*, on the other hand, plays absolutely no part in either case, even though Carston claims that it is encoded by 'flaunt'. Its role in Carston's account of the interpretation process is almost as insignificant as its role in her account of language (mis)use by speakers S and M. The hearer S or M does decode the utterance and access the lexical concept, but replaces it with the new and

unrelated ad-hoc concept *flout* once he decides that *flaunt* is far from relevant in the circumstances. So all in all the lexical concept supposedly encoded by 'flaunt', *flaunt*, has just one role in the entire communication process, and a rather negative one at that: it has to be rejected, and completely dropped (as opposed to being loosened, and partially dropped) in order for the hearer to derive relevant contextual effects.

All of the above follows from Carston's view of E-physical utterance as fundamental to language use, according to which the SEMANTIC properties of a linguistic expression could be seen as somehow secondary to the PHONETIC properties. (Or, in Saussurean terms, the signified could be seen as secondary to the signifier.) This concept of language use is shared by Sperber and Wilson, by Chomsky -- and, in theory, by anyone else who subscribes to the idea of linguistic expressions as vehicle-meaning pairs. As the "VEHICLE" for MEANING, the *E-physical phonetic properties* are generally seen as the most salient elements of a linguistic expression when it is *E-physically uttered* -- so much so that the linguistic expression is represented as being used even when what it supposedly encodes is disregarded by the speaker and discarded by the hearer. In fact it is when this happens that the fundamental issue at stake in all this becomes particularly clear. The *use* of a linguistic expression, as conceived by those who regard linguistic expressions as vehicle-meaning pairs, could be described as the production of the E-physical vehicle to express its meaning, or the expressing of the meaning through its vehicle. The *misuse* of a linguistic expression, on the other hand, is always seen as a matter of producing the phonetic vehicle to express the "wrong", unencoded meaning (e.g. misusing 'flaunt' by uttering [flɔ:nt] to express *flout*), and never of expressing the semantic meaning through the "wrong", unencoding vehicle (e.g. misusing 'flout' by expressing *flout* with the phonetic vehicle [flɔ:nt]). This implies that the identity of the linguistic expression apparently uttered is determined by, and only by, the phonetic form produced.

Thus Carston claims that the speaker who attempts to communicate that Mary belongs to the upper classes by uttering

(1) Mary is one of the hoi polloi.

has used the linguistic expression 'hoi polloi', but misused it to express the "wrong" meaning (i.e. a meaning which the phonetic form of 'hoi polloi' does not encode). In other words, Carston assumes that 'hoi polloi', encoding *common people*, is the linguistic expression which the speaker has uttered, and which the hearer perceives and maps

> ... "correctly" to a conceptual address for [common folk]... (ibid: 17)

before (presumably) going on to construct the ad hoc concept *the upper classes*. But Carston also mentions the possibility of a hearer who is

> ... also mistaken, and in the same way as the speaker, in his understanding of the expression 'hoi polloi'... (ibid)

Here, Carston appears to be saying that the linguistic expression is *misused* by the second hearer as well the speaker -- both of them are "mistaken" about 'hoi polloi', the speaker misusing it to mean, and the hearer misinterpreting it as meaning, *the upper classes*. The implication, of course, is that 'hoi polloi' is specified as encoding *common people* by a definitive set of linguistic rules which both speaker and second hearer deviate from in their (mis)use of the expression. But there is no denying the fact that what is an ad-hoc concept for the first hearer is recovered by the second hearer via linguistic decoding; and that, as far as cognitive effects are concerned, the speaker uses a different linguistic encoding for each hearer: an encoding of *common people* for the first hearer, an encoding of *upper classes* for the second.

Before I continue, it is necessary to distinguish *hearer* M from *hearer* S. The second hearer referred to in the previous paragraph is, of course, hearer M -- like speaker M, his

system of linguistic knowledge does not include the encoding of *common people* by 'hoi polloi'. There is also hearer S, who does believe, for example, that 'flout' encodes *flout*, but who recovers *flaunt* from an utterance of 'flout' as a result of a "perceptual error". The case of speaker and hearer S corresponds to Chomsky's account of "individual misuse" (see section 2.3), which he claims is the only sense in which language (i.e. I-language) is misused. Speaker S speaks in a way "that is not in accord with his I-language" (Chomsky 1992a: 121), while hearer S "[offers] judgements inconsistent with his I-language" (ibid). But if speaker or hearer *M* is mistaken it is certainly not in virtue of deviating from "HIS" internal language or linguistic code (to use Chomsky's possessive pronoun) -- it would be because his entire system of linguistic knowledge has deviated from some other (perhaps definitive) code which he does *not* possess, and whose specifications about what particular utterances encode have absolutely no effect on how he uses language. Now, even though it is not a part of his linguistic code, Carston might assume (wrongly, in my opinion) that speaker M misuses the encoding of *common people* by 'hoi polloi' by voluntarily producing a phonetic form which a certain linguistic rule matches with *common people* -- i.e. by *uttering* the linguistic expression 'hoi polloi' -- to express the "wrong", unencoded *upper classes*. But hearer M, on the other hand, does nothing other than automatically decode the utterance to recover *the upper classes*. So for hearer M, at least, the speaker has used the encoding of *the upper classes* by 'hoi polloi', and used it correctly.

Or rather, *"correctly"*, with scare quotes around the word. "Correctly" is how Carston presents it -- not in respect of hearer M, it has to be stressed, but the hearer who decodes 'hoi polloi' "correctly" to recover *common people* (see above, and ibid: 17). This implies that the notions of use and misuse are relative -- and, in this case, that whether the speaker "misuses" 'hoi polloi' or perhaps uses it "correctly" depends on what concept(s) the hearer and the speaker map the utterance to. It follows from this, in my view, that to claim that a speaker misuses a linguistic expression is to beg the question, for we must

first establish what linguistic expression is actually being used before we can say that it is being misused. Carston all but concedes this in a footnote:

> Of course, the matter of what gets called an error or a misuse is sometimes more a matter of power and authority than anything to do with language itself... (ibid: 82, fn 2)

She also mentions "disparate encodings" (ibid) -- i.e. the fact that different people may have different 'systems of linguistic knowledge', different beliefs about what words mean. This suggests that the *code* itself is ad-hoc in nature; and that the issue of what is used or misused has nothing to do with what is uttered or what is determined by some supposedly fixed and autonomous set of linguistic rules, but instead involves the disparate codes of different language users. For instance, nothing would be misused when speaker M utters (1) to hearer M: as far as both their linguistic codes are concerned, speaker M has provided the best linguistic evidence for his intentions. But if speaker S utters (1) to hearer S, each would consider the other as having misused 'hoi polloi'. And in M/ S pairings only speakers or hearers S would think that the linguistic expression has been misused.

To conclude this section, I have two points to make. Firstly, I am not arguing that a cognitive theory like Carston's or Sperber and Wilson's is unable to account for phenomena like misuse, slips and disparate encodings -- only that first principles about language and language use need to be clarified, and set out in far greater detail. Secondly, note once again that all the cases of misuse listed above involve the production of a phonetic form to express a concept which it is not supposed to encode, and that it is the phonetic form and not the concept which is assumed to determine what linguistic expression is misused. The reverse appears to be true for Davidson. Whereas *flout* for Carston is the "wrong" concept, not encoded by what is uttered by speaker M or S, Davidson would consider it to be the *literal meaning* of what the speaker uttered. Davidson's 1986 discussion of malaprops is the subject of the next section.

4.2 Davidson's prior and passing theories

Carston's pragmatic account of misuse and slips is underpinned by the notion of public, definitive sets of linguistic rules (i.e. linguistic codes) which determine the literal meanings of the phonetic vehicles produced/ uttered by speakers, regardless of the individual speaker's intentions. Davidson, on the other hand, claims that the phenomenon under investigation involves the taking over, not only of literal meaning by the speaker's intended meaning, but also of one system of linguistic rules by another. In his 1986 paper "A nice derangement of epitaphs" (henceforth "Derangement"), an analysis of malaprops -- which he places in the same category as "incomplete or grammatically garbled" utterances, slips of the tongue, words that we have never heard before and new idiolects (1986: 437) -- leads Davidson to the conclusion that "there is no such thing as language" (ibid: 446). This is not as melodramatic as it sounds, especially with the qualification that follows: "... not if a language is anything like what many philosophers and linguists have supposed" (ibid). Clearly the aim of "Derangement" is not to show that language (whatever it is) does not exist, but to demonstrate that phenomena like malaprops undermine the theories of "many philosophers and linguists" about language and linguistic meaning.

It has to be stressed that Davidson begins by affirming that

> ... nothing should be allowed to obliterate or even blur the distinction between speaker's meaning and literal meaning. (ibid: 434)

He also has a way of defining literal meaning, which he calls "first meaning", in terms of what speakers or interpreters do. First meaning could be said to come first in the order of interpretation, except that

> ... it often happens that we can descry the literal meaning of a word or phrase by first appreciating what the speaker is getting at. (ibid: 435)

Alternatively, the first meaning of a linguistic expression which a speaker utters may be located through the speaker's intentions in uttering the linguistic expression, and how these intentions are

> ... usually unambiguously ordered by the relation of means to ends (where this relation may or may not be causal)... (ibid)

For instance, Peter utters the LE 'Eat your eggplant' (example from Davidson 1981: 271[6]) with the intention of uttering an expression that means *eat your eggplant*, with the intention of telling Mary to eat her eggplant, with the intention of getting her to eat her eggplant. Peter also has "the Gricean intentions to achieve certain of these ends through [Mary's] recognition of some of the intentions involved" (1986: 435). The first meaning of 'Eat your eggplant' is specified by the first intention which requires this self-referring Gricean feature.

Note that Davidson's "first meaning" is NOT meaning endowed by the speaker's intentions, but merely, though necessarily, intended by the speaker to be grasped -- and grasped *first* -- by the hearer. To be more precise, it is meaning which the speaker necessarily wishes the hearer to grasp or register (and I deliberately use these vague verbs, rather than "interpret") in virtue of having understood the meaning of the expression uttered, but which does not necessarily constitute the ultimate and complete interpretation which he intends his utterance to receive. Note also that first meaning is defined here in terms of nothing more than the SPEAKER'S intentions -- intentions which the HEARER may *not* comply with. That is, what the speaker intends the hearer to grasp first need not

---

[6] The page numbers for Davidson 1981 are take from the 1984 reprint.

necessarily be the first in the order of the hearer's interpretation (hence Davidson's rejection of the "first in the order of interpretation" definition).

Davidson goes on to narrow down this notion of first meaning -- which, like Grice's non-natural meaning, "applies to any sign or signal with an intended interpretation" (ibid: 436) -- to one of *linguistic* meaning, by applying to it three "plausible" principles concerning first meaning in language (ibid: 436-7).

The first principle requires that first meaning be (1) SYSTEMATIC; or rather, that there should be *systematic relations* between the meanings -- the first meanings -- of utterances. These systematic relations would give rise to semantic and structural properties, through which "a competent speaker or interpreter" (ibid: 436) interprets his own or others' utterances. Davidson also argues that this system of interpretation "has a finite base and is recursive" (ibid: 438); and proceeds to call it a *theory*, "as if the interpreter were using the theory we use to describe his competence" (ibid).

The second principle states that first meaning must be (2) SHARED. I.e. the system or theory of first meanings must be shared by speaker and interpreter in such a way that the interpreter uses it "to understand the speaker", and the speaker, "to guide his speech" according to "how the interpreter will interpret him" (ibid: 437).

The third principle states that first meaning must be (3) PREPARED: "learned in advance of occasions of interpretation and... conventional in character" (ibid: 436).

Note that both Davidson's notion of first meaning and his three principles about linguistic meaning presuppose that the sort of meaning which he is concerned with is *for conveying and interpreting*. A system or theory of such meanings -- i.e. a language -- is therefore a *theory of interpretation*. Principle (2), in particular, does not seem so much a principle of language as one of communication. Davidson claims that the notion of shared systems of first meanings "does not demand that speaker and interpreter speak the same language" (ibid: 437) -- only that the interpreter shares the speaker's understanding of the *speaker's* words. Thus principle (2) in effect states that, whatever theory of interpretation

is used (and speaker and interpreter may have different theories), what is required for successful communication is that the interpreter manages to interpret the speaker's words or signals as the speaker understands them.

One might argue that it is this view of language as a theory of interpretation, and not the phenomenon of malapropisms, which threatens the distinction between literal and speaker's meanings. But, as I have mentioned, Davidson does distinguish between (a) interpretative processes which count as one's "basic linguistic competence" and (b) those which are governed by general, pragmatic principles like the Gricean Maxims of Conversation. In other words, he does distinguish between (a) the literal or first meanings of LEs and (b) what is implied or implicated by someone who uses them. He stresses that the problems created by malapropisms and the like are quite different from problems concerning the (a)-(b) distinction. In the case of metaphor, for example, the question would be how a particular linguistic expression whose literal meaning the language user is fully aware of, and has deliberately chosen (together with the other elements of the linguistic expression), is used to convey some other meaning. Or, in Davidson's words, phenomena like metaphor or irony are "achieved by way of the normal meanings of the words" (ibid: 439).

What Davidson is interested in is what Mrs Malaprop does in *The Rivals*, and what Donnellan does in his (1968) reply to MacKay's (1968) discussion of Donnellan's (1966) distinction between the referential and attributive uses of definite descriptions -- and the fact that

> ... in all these cases the hearer has no trouble understanding the speaker in the way the speaker intends. (1986: 434)

The main details of the Donnellan case are as follows. Donnellan claims that someone who says 'Smith's murderer is insane' has said something true as long as he is using 'Smith's

murderer' *referentially*, to refer to a particular person who is insane -- even if that person did not murder Smith. MacKay's objection is that this amounts to Humpty Dumpty's theory of meaning, whereby one uses an expression (e.g. 'Smith's murderer') to mean whatever one wants it to mean (e.g. *the insane man who did not murder Smith*). For Davidson, however, Donnellan's referential use has "nothing to do with words changing their meaning or reference" (Davidson 1986: 439). Even if the speaker uses 'Smith's murderer' to refer to someone who did not murder Smith,

> ... the reference is none the less achieved by way of the normal meanings of the words. (ibid)

Once again Davidson insists on a clear distinction between literal meaning and the speaker's intended meaning, arguing that, while the speaker may have expressed something that is true, he has done so by using a sentence that is literally false.

Davidson, of course, is much more interested in the speaker who does *not* proceed "by way of the normal meanings of the words" when he uses those words to mean something else. Humpty Dumpty, Davidson maintains, is NOT such a speaker, because he cannot be said to *mean anything* when he tells Alice 'There's glory for you':

> He cannot mean what he says he means because he knows that 'There's glory for you' cannot be interpreted by Alice as meaning 'There's a nice knockdown argument for you.' We know he knows this because Alice says 'I don't know what you mean by "glory" ', and Humpty Dumpty retorts, 'Of course you don't -- til I tell you.' (ibid: 440)

Davidson takes Donnellan's view that

> A speaker cannot... intend to mean something by what he says unless he believes his audience will interpret his words as he intends (the Gricean circle). (ibid: 439)

Thus Davidson claims that one can change the meaning of an expression as one intends,

> ... provided you believe (and perhaps are justified in believing) that the interpreter has adequate clues for the new interpretation. You may deliberately provide those clues... (ibid)

An example of the speaker himself voluntarily providing such clues is when Donnellan ends his reply to MacKay with the expression 'There's glory for you'. By that Donnellan means *There's a nice knockdown argument for you*. Unlike Humpty Dumpty, however, he *IS* interpreted as meaning *There's a nice knockdown argument for you* -- not only because he believes that the reader would interpret it to mean what he intended it to mean, but also because he uses it in a context which he knows would prompt the reader to interpret it as *There's a nice knockdown argument for you*.

Davidson refers to Donnellan's 'There's glory for you', and Mrs Malaprop's 'a nice derangement of epitaphs', as cases of the speaker "getting away with it" (ibid:440). What "getting away with it" involves is this:

> ... the interpreter comes to the occasion of utterance armed with a theory that tells him (or so he believes) what an arbitrary utterance of the speaker means. The speaker then says something with the intention that it will be interpreted in a certain way, and the expectation that it will be so interpreted. In fact this way is not provided for by the interpreter's theory. But the speaker is nevertheless understood; the interpreter adjusts his theory so that it yields the speaker's intended interpretation. The speaker has 'gotten away with it'. (ibid)

Furthermore,

> The speaker may [like Donnellan] or may not [like Mrs Malaprop]... know that he has got away with anything; the interpreter may [like Donnellan's reader] or may not [like Mrs Malaprop's interpreter] know that the speaker intended to get away with anything. (ibid)

To account for speakers "getting away with it", Davidson proposes that at every stage of what he calls a "speech transaction" speaker and interpreter have not one but two theories of interpretation: a PRIOR THEORY and a PASSING THEORY. During the speech transaction both theories are continually revised by both speaker and interpreter. The prior theory for the interpreter is "how he is prepared in advance to interpret an utterance of the speaker", and, for the speaker, "what he *believes* the interpreter's prior theory to be". The passing theory for the interpreter is "how he *does* interpret the utterance", and, for the speaker, "the theory he *intends* the interpreter to use" (ibid: 442).

Before the speaker starts speaking, the interpreter's prior theory is adjusted according to whatever relevant information about the speaker and the speaker's idiolect that he is able to take into account: information about

> ...the character, dress, role, sex, of the speaker, and whatever else has been gained by observing the speaker's behaviour, linguistic or otherwise. (ibid: 441)

During the course of the utterance the adjustment continues, and the prior theory continues to change "in the light of new evidence" (ibid), some of it provided by the speaker himself. But once he starts to interpret the utterance, there is every possibility that the theory which he actually uses, the passing theory, is not the theory he was prepared to use, the prior theory. (Like the prior theory, however, the passing theory keeps changing during the utterance.)

The speaker, for his part, has a prior theory which takes into account information about the interpreter and the interpreter's prior theory. The speaker is aware of the advantage of uttering expressions which "he believes can and will be interpreted in a certain way" (ibid: 442). But when it comes to the point, he may not produce an utterance in accordance with what he believes is the interpreter's current theory -- he may

"deliberately dispose the interpreter to modify [it]" (ibid) to converge with his (the speaker's) own passing theory.

The problem is that neither the prior nor the passing theory of interpretation conforms to all three of Davidson's principles (1)-(3) concerning the nature of linguistic first meaning. Both theories, being full recursive theories, satisfy (1) -- even though the passing theory's "expected field of application is vanishingly small" (ibid). A passing theory must be shared (principle (2)) in the sense that communication is successful only if the theory which the speaker intends the interpreter to use coincides with the theory which the interpreter actually uses to interpret the speaker's utterance (ibid).[7] But it is certainly not "learned in advance of occasions of interpretation" (principle (3)), being constructed for, and applied to, a particular utterance on a particular occasion, and containing "its changing list of proper names and gerrymandered vocabulary", and "every successful -- i.e. correctly interpreted -- use of any other word or phrase, no matter how far out of the ordinary" (ibid).

A prior theory, on the other hand, might be said to be "prepared" -- but only up to a point. A passing theory would, of course, be "geared to the occasion" (ibid: 441). However, as I have mentioned, even a prior theory takes into account information about the interlocutor and his idiolect. Thus interpreters "must be expected to have quite different prior theories for different speakers", and speakers

> ... always have the interpreter in mind; there is no such thing as how we expect to be interpreted. We inhibit our higher vocabulary, or encourage it, depending on the most general considerations, and we cannot fail to have premonitions as to which of the proper names we know are apt to be correctly understood. (ibid:443).

---

[7] Davidson adds: "Of course, there are degrees of success in communication..." (1986: 442). Since humans are non-telepathic, it is possible that completely identical passing theories of speaker and interpreter, and therefore complete success in communication, is never achieved.

It also follows from this that the prior theories of speaker and interpreters are neither shared nor even assumed to be shared. In fact Davidson maintains that there is no reason why they should be (ibid), precisely because shared prior theories are not necessary for successful communication. He cites the example of Mrs Malaprop and her interpreter: it does not matter that Mrs Malaprop's *prior theory* about 'a nice derangement of epitaphs' is that it means *a nice arrangement of epithets* while the interpreter has a *prior theory* according to which 'a nice derangement of epitaphs' means *a nice derangement of epitaphs*, for the interpreter would understand Mrs Malaprop as long as their *passing theories* coincide.

So the argument in "Derangement" goes like this. Davidson begins by asserting (ibid: 437, 440-1, 443) that malaprops do not simply involve the use of linguistic expressions to mean something other than what they mean. He subscribes to the view that, no matter how remote the speaker's intended meaning is from the literal meaning of the linguistic expression, the linguistic expression can be said to be used only if it is -- at least partly -- used in virtue of its literal meaning. The speaker who utters a malaprop, on the other hand, appears to use the linguistic expression in such a way that the new meaning expressed IS the first meaning that he intends the interpreter to access. To account for the interpreter's ability to register and apply the change in literal meaning, Davidson suggests that he has two simultaneous theories of interpretation. The prior theory contains the old meaning; it is the passing theory which accommodates the new meaning as literal linguistic meaning:

> Every deviation from ordinary usage, as long as it is agreed on for the moment (knowingly deviant, or not, on one, or both, sides), is in the passing theory as a feature of what the words mean on that occasion. Such meanings, transient though they may be, are literal; they are what I have called first meanings. (ibid: 442-3)

But neither prior nor passing theory conforms to principles (1)-(3). The passing theory is obviously not learned in advance, while the prior theory does not have to be (and

usually is not) shared -- thus neither theory can be described as being governed by antecedently shared conventions. From this Davidson concludes that there is no common and fixed set of linguistic conventions which determines what utterances mean in particular acts of communication and interpretation. Instead, speakers and interpreters construct convergent passing theories on the basis of

> ... wit, luck, and wisdom from a private vocabulary and grammar, knowledge of the ways people get their point across, and rules of thumb for figuring out what deviations from the dictionary are most likely. (ibid: 446)

4.3 Dummett's reply to Davidson: language as social convention

Davidson's arguments and conclusions in "Derangement" are aimed at the sort of concept of language that Dummett subscribes to. In his (1986) reply to "Derangement", Dummett objects to Davidson's notion of INTERPRETATION and the very central place it occupies in the "Derangement" account of language and communication. As Dummett points out, Davidson's language users are always speakers and *interpreters* rather than speakers and *hearers*. Theoretically, Dummett's conception of language as a "social practice" could not be more opposed to the views of those who believe that social practices belong to the domain of the extralinguistic. Nevertheless, he joins them in claiming -- though of course on grounds supposedly completely different from internalist/ intensionalist ones -- that there is a clear distinction between interpretation, on the one hand, and linguistic parsing or decoding, on the other; and, co-extensively, between utterance/ speaker's meaning and literal linguistic meaning.

Dummett cites (ibid: 464, 467) and follows Wittgenstein's distinction between *Deutung* and *Auffassung* (*Investigations*, 201). For Dummett, a *Deutung* is an

interpretation "in the strict sense" (Dummett 1986: 467) -- when, for instance, a hearer has to "search for the speaker's meaning" (ibid: 464) -- and an *Auffassung*, a way of understanding or grasping which excludes any understanding of *how* what is understood is actually understood. He observes that Davidson's theory of language is based exclusively on the former, the "exceptional cases", and completely ignores the latter, the "normal case" (ibid: 471). In the "normal case", Dummett argues, the speaker or hearer's use of language does not consist in "having any beliefs about the other person" -- for example, having beliefs or expectations about the other person's prior theories. The only beliefs involved ("if they can be called beliefs" (ibid: 472), Dummett parenthesises) are "about what the words mean, not about what the other takes or intends them to mean" (ibid). Speakers and hearers just "go on these beliefs" (ibid):

> In the normal case, the speaker simply says what he means. By this I do not mean that he first has the thought and then puts it into words, but that, knowing the language, he simply speaks. In the normal case, likewise, the hearer simply understands. I.e., knowing the language, he hears and thereby understands; given that he knows the language, there is nothing that his understanding the words consists in save his hearing them. (ibid: 471)

Speakers and hearers do more than this, engaging in *interpretation*, only in the "exceptional cases" where "there are indeed no [linguistic] rules to follow" -- when the hearer "has to apprehend the speaker's intention much as he has to apprehend the intentions behind non-linguistic actions" (ibid: 474), or when the speaker "wants to communicate and is conscious of obstacles to understanding" (ibid: 473).

Dummett's "normal" and "exceptional" cases must also be considered within the framework of his own theory of language and language use. While Davidson equates languages with theories of interpretation, Dummett not only distinguishes interpretation from "normal" linguistic behaviour, but also the latter from language itself and its

semantics. A LANGUAGE, for him, is a *phenomenon*, consisting of a set of expressions with structure and meanings (ibid: 467). What it is for those expressions to have the meanings that they have is made explicit by a FIRST-ORDER THEORY OF MEANING for the language. This theory of meaning "explains" speech (and other, probably secondary forms of linguistic behaviour to do with literacy) -- but does so only *indirectly*, in that it contains no reference to anything language users do or think (ibid). Thus LINKING PRINCIPLES are required to connect it to the thoughts and actions of the person who speaks the language, and who thereby has a SECOND-ORDER THEORY about the first-order theory of meaning (second-order in virtue of being a theory about another theory (ibid: 466)). Dummett maintains that "no speaker knows every word of the language or uses correctly every word he does know" (ibid: 469), and regards the second-order theory as an individual's "imperfect" (ibid: 475), "partial, and partly erroneous" (ibid: 468) grasp of the first-order theory of meaning of the language. Consequently, the linking principles which connect it to the perfect, complete first-order theory of meaning have to involve the notion of an idiolect (ibid: 469) -- or, to be more precise, the notion of an idiolect as the individual speaker-hearer's imperfect, incomplete second-order theory about the perfect, complete first-order theory of meaning for the "common language".

However, Dummett stresses that this second-order theory is NOT a speaker-hearer's *interpretation* of the first-order theory. In fact it operates like a first-order theory --

> Its literal content is that certain expressions have certain meanings, or that some individual intends them to have or takes them as having those meanings. (ibid: 468)

-- though unlike the first-order theory it does not state what it is for those expressions to have the meanings that they have.

Dummett argues that

> ... a common language is related to an idiolect essentially as the rules of a game are related to a player's beliefs about what they are. (ibid: 469)

He compares first-order theories of linguistic meaning with (representations of) the rules of games, speaker-hearers with players of a game, and idiolects with players' grasp of the rules of a game (ibid: 469, 472, 473). The important point here, in respect of Davidson and "Derangement", is that

> ... the players' grasp of the rules does not consist in any theories they have about the knowledge of the rules on the part of the other players. (ibid: 472)

-- they simply take for granted that the other players know the rules. Similarly, speakers and hearers in the "normal case" take for granted that they have a common language (ibid: 473).

For Davidson, apparently, there is no such "normal case". Dummett's "exceptional cases" *are* Davidson's "normal case", and the theories used in Davidson's "normal case" are also what Davidson regards as languages. Thus, on Dummett's terms, a language for Davidson's speaker or *interpreter* (rather than *hearer*) is not just a second-order theory but one whose subject is *another second-order theory*, that of the other speaker or interpreter. In other words, it constitutes an *interpretation* of the other speaker or interpreter's (also second-order) theory of interpretation, and is the basis upon which one might construct an *interpretation* of the actions or behaviour determined by that other theory of interpretation.

Dummett does not disagree with Davidson's point that we constantly adjust our theories of interpretation according to circumstances, and according to who our interlocutors are in particular. His main objection, as I have mentioned, is Davidson's equation of language with theories of interpretation. He maintains that Davidson's theories of interpretation -- whether passing or prior -- are *all* second-order theories. This is

because of the "quasi-Gricean" nature of Davidson's account of linguistic meaning -- i.e. because it appears to be wholly based on speakers/ interpreters' expectations of, and intentions towards, one another. The resulting absence of first-order theories of meaning, Dummett argues, means that all Davidson's theories must be of infinite order -- and that his theory of language is an "exitless loop" of speaker's theories about interpreter's theories, and interpreter's theories about speaker's theories about interpreter's theories, and so on (ibid: 470). Thus Dummett regards Davidson's theory of language as "a more sophisticated version" of Humpty Dumpty's idea of linguistic expressions meaning whatever the speaker wants them to mean, and a refutation of Alice's idea of linguistic expressions having meanings independently of language users (ibid).

According to Dummett, Davidson arrives at this "Humpty Dumpty" view of language from "[concentrating] too exclusively on communication" (ibid: 471). While he believes that the primary role of language is as an instrument of communication, and that "language is a vehicle of thought because it is an instrument of communication, and not conversely" (ibid: 470-1), Dummett also maintains that one cannot build a theory of language around the phenomenon of verbal communication alone, if only because language is supposed to have that other (though secondary) role as a vehicle of thought which the theory must also be able to account for. Dummett cites Wittgenstein again, this time appealing to his challenge to say one thing and mean something else (ibid: 471; *Investigations*, 510), and arguing that the "difficulty" of saying 'The sky is clouding over' and meaning 'There is no odd perfect number' has nothing to do with failing to meet the interpreter's expectations, since it is "just as great if you are saying it to yourself" (ibid: 471). As Dummett sees it,

> ... the difficulty lies in the fact that 'The sky is clouding over' does not mean 'There is no odd perfect number'. (ibid)

-- i.e. in the fact that the linguistic expression 'The sky is clouding over' means *The sky is clouding over* and not *There is no odd perfect number*, regardless of the fact that the speaker or thinker intended to say or think 'The sky is clouding over' and mean *There is no odd perfect number*.[8]

Dummett, of course, believes that linguistic expressions have "standard meanings", and that no process of interpretation can explain what it is for the hearer or speaker to attach a standard meaning to the corresponding linguistic expression. Note that these "standard meanings" are not assigned to Dummett's linguistic expressions by a Chomskian I-grammar, or by some Fodorean (or relevance-theoretic) input system upon which utterances automatically impinge, but by the rules or conventions of a SOCIAL PRACTICE. Language for Dummett is "a practice in which people engage" (ibid: 473), and a practice is essentially social in that it is "learned from others and is contituted by rules which it is part of social custom to follow" (ibid). Being social in nature, linguistic

---

[8] According to Dummett, even the involvement of both speaker and hearer (as opposed to the obviously solitary thinker) in a case of language use does not guarantee that there will be the sort of reciprocal theory-construction upon which Davidson's theory of language is based. This, he claims, is because it is "possible to mean something by an expression without intending one's hearer to understand it" (1986: 472). Dummett gives two examples of this, one where the speaker does not address a particular hearer or particular hearers, and one where he does. The first is the soliloquizing speaker, who has no intention of being understood by the hearer, "a mere eavesdropper" (ibid). The second is the Eric Blore character in the film *Top Hat*, who utters insults in English to the face of an apparently non-English-speaking policeman and -- more importantly -- thinks that he is actually performing the act of insulting the policeman to his face even though he believes that his addressee does not understand a word he is saying. In my opinion, neither of these cases can be regarded as direct counterexamples to . claims about *communication* being the basis of theories of linguistic meaning. This is because the speaker-hearer relations in both examples appear to be essentially different from those which Davidson considers in "Derangement". With regard to the first example, the fact that the eavesdropper is not being addressed by the speaker clearly puts him in a different category from Davidson's interpreters. In fact it is possible that the first example may be taken as undermining Davidson's case only insofar as it involves a speaker who is in effect speaking to himself. So, as a counterexample to "Derangement", it is more closely related to Wittgenstein's experiment about saying one thing and meaning another. The speaker does address the hearer in the second example, but with what appear to be twofold and contradictory intentions -- to insult the policeman to his face, and to *not* be understood (and therefore to escape arrest). One might argue that the latter intention cancels out the former -- which, of the two, is the truly communicatory intention -- on the grounds that it is impossible to insult someone to his face without being understood by him.

expressions do not have meanings "intrinsically, and hence independently of anything human beings do" (ibid). They are, however, described by Dummett as having meanings "in themselves", and "independently of any particular speakers" (ibid). This means that the language user simply uses a linguistic expression to mean what it means, as specified by the conventions of the language to which it belongs, without having to form any express intention or theory about other language users, the context of use, or the linguistic expression/ language itself. Of course he may want to use the linguistic expression to mean something else, in which -- "exceptional" -- case the process of interpretation or second-order theory construction comes into play. But on the whole language users have a "responsibility to the language as a social institution", and one cannot theoretically "liberate" them from this "responsibility" as Dummett thinks Davidson does, by giving their intentions and expectations precedence over linguistic conventions (ibid).

For Dummett, the concept of CONVENTION is fundamental to a theory of language -- conventions being "what constitute a social practice" (ibid), and a social practice being what Dummett regards language to be. This linguistic role of convention, he maintains, is repudiated by the Humpty Dumpty view of language, including Davidson's "refined" version. At the end of "Derangement" Davidson asserts that there is "no such thing as a language" -- i.e. language as "shared conventions, rules, or regularities" -- "to be learned, mastered, or born with" (Davidson 1986: 446). He concludes that convention is either related in some other way to language, or perhaps not at all. But at the same time, as Dummett observes, Davidson claims that

> ... we may know so little about our intended interpreter that we can do no better than to assume that he will interpret our speech along what we take to be *standard lines*. (ibid: 443, my italics)

-- an assumption which, in its turn, is based on the assumption that the interpreter "belongs to our *language community*" (ibid, my italics). To Dummett these references to

"standard" methods of interpretation corresponding to particular "language communities" indicate that the notion of linguistic conventions, of speaking the same language, is absolutely necessary even for Davidson's interpretation-based theory of language. What we bring to a successful linguistic encounter with another -- Davidson's prior theory, perhaps -- must be something like a common set of rules governing what particular utterances mean; otherwise nobody would know what any utterance means, let alone what the speaker intended to convey in uttering it. In short, Dummett's argument is that there must be something which the speaker and hearer share, without which any of their utterances could mean anything to either of them.

4.4 Language use, literalness and convention

Davidson's conclusion that knowing a language is the same as "knowing our way around in the world generally" (1986: 446) follows directly from his view of language as a theory of *interpretation*. Such a view of language is obviously unacceptable to those who distinguish between (i) language as constituted by or involving a fixed and autonomous set of rules or principles; and (ii) the *uses* to which language may be put, involving, as Chomsky puts it, "all other capacities of the mind" (1992a: 120). According to this distinction, utterance interpretation does not constitute the linguistic, merely the *use* of the linguistic; and it is not *language* that changes to accommodate relevant contextual information, as Davidson appears to believe, but what the language user does with language that has to be adjusted. Thus, for Chomsky, language use is the business of performance systems, "interfacing" with "other capacities of the mind". For Carston and the relevance theorists, it is the universal central systems that carry out the cognitive processes of utterance interpretation. For Dummett, language users have second-order

theories about linguistic meaning, as well as about other language users' second-order theories about linguistic meaning.

In formulating his passing and prior theories, on the other hand, Davidson appears to focus exclusively on communication, on utterance production and interpretation, and in doing so could be regarded as needlessly conflating the truly linguistic elements with features of how speakers and hearers USE language in respect of one another. Both prior and passing theories require at least some constructing on the part of speaker or hearer, taking into account certain aspects of the utterance context (especially information about the interlocutor) and incorporating what would be *contextual implications* for relevance theorists. Moreover, Dummett adds that the passing theories which interpreters construct for specific utterances are not really theories at all: they

> ... [bear] only on those utterances for which the [prior theory] does not yield the correct interpretation. (1986: 466)

and

> ... so viewed, will not be a structured theory, but only a collection of disconnected propositions. (ibid)

It is certainly true that Davidson sees language very much as a matter of performance rather than competence. This comes across particularly clearly on one of the rare occasions in "Derangement" that he actually talks about the speaker's *use* of linguistic expressions. Within the "Derangement" framework, Mrs Malaprop's 'a nice derangement of epitaphs' or Donnellan's 'There's glory for you' are expressions belonging to a new, passing theory of the interpreter (and, in Donnellan's case, the speaker as well) whose prior theory states that 'a nice derangement of epitaphs' means *a nice derangement of epitaphs*, not *a nice arrangement of epithets* (or that 'There's glory for you' means *there's*

glory for you, not *there's a nice knockdown argument for you*). There is no doubt that Davidson believes this. However, with reference to the two examples, he also remarks:

> There is no word or construction that cannot be converted to a NEW USE by an ingenious [i.e. Donnellan] or ignorant [i.e. Mrs Malaprop] speaker. (1986: 441, my caps)

In view of his claim that Donnellan's 'There's glory for you' and Mrs Malaprop's 'A nice derangement of epitaphs' belong to new theories of interpretation, Davidson seems to be saying here that the new (passing) theory of interpretation is created by the new *use* of the expressions of the old (prior) theory.

Even his discussions of first, literal, conventional linguistic meaning are based on what language users do with that meaning -- in other words, on the intentions and expectations of speakers and hearers in respect of that meaning. In "Derangement" he defines first/ literal meaning as that which the speaker *intends* the hearer to access first. In his earlier (1981) paper "Communication and convention" he relates the notion of literal linguistic meaning to Lewis' 1975 analysis of convention. Lewis defines a convention as a regularity R in which more than one person must be involved. The properties of R are as follows:

> (1) Everyone involved conforms to R and (2) believes that others also conform. (3) The belief that others conform to R gives all involved a good reason to conform to R. (4) All concerned prefer that there should be conformity to R. (5) R is not the only possible regularity meeting the last two conditions. (6) Finally, everyone involved knows (1)-(5) and knows that everyone else knows (1)-(5), etc. (Davidson 1981: 276, after Lewis 1975: 5-6)

This, too, is in terms of the relation between convention and "everyone involved" -- it has to do with people conforming to R, rather than the intrinsic nature of R itself.

Futhermore, the intentions and expectations of speakers/ hearers are as central to Davidson's application of Lewis' R to language, as they are to the relevance theorists'

account of language use in ostensive-inferential verbal communication. For speaker and hearer to conform to Davidson's version of a linguistic convention, both must understand the speaker's words. Also,

> ... [(A)] the speaker must intend the hearer to interpret his words in the way the speaker intends, and he must have adequate reason to believe that the hearer will succeed in interpreting him as he intends. [(B)] Both speaker and hearer must believe that the speaker speaks with this intention, and so forth; in short, many of Lewis' conditions would seem to be satisfied. (1981: 277)

Note that what I have labelled (A) is similar to the conditions of optimal relevance, while (B) seems to correspond to the principle of relevance (i.e. that the presumption of optimal relevance, or (A), is ostensively communicated, made mutually manifest). Therefore, what bears upon language use and the central cognitive processes for the relevance theorists is for Davidson a matter of linguistic convention (or, in terms of relevance theory, linguistic encoding and decoding).

However, there may be more to Davidson's arguments than an apparent externalist confusion of language with language use, of the semantic with the pragmatic. Firstly, Davidson 1981 concurs with Black 1972/3 and Chomsky 1971 that communication is not the only activity that involves language. Black observes that

> ... a man may outline a lecture, or write a note to remind himself of an appointment, or simply utter words... in the absence of an audience. (1972/3: 264)

Chomsky claims that the use of language "need not involve communication, or even the attempt to communicate" (1971: 19). These last two quotes bear upon Black's and Chomsky's *denial* that

> ... one must always intend to produce some non-linguistic effect through having one's words interpreted. (Davidson 1981: 272)

Davidson in turn argues that intended (non-linguistic) effects need not necessarily "involve someone other than the speaker" (ibid: 273). The important point is that they are supposed to arise in some way or other from the literal meanings of the linguistic expressions which the speaker (soliloquizing or otherwise), writer or thinker has deliberately chosen, over other words with their other meanings, to use.

Secondly, Davidson does appear to believe that these intended effects -- or "ulterior purposes", as he calls them (1981: 272) -- arise from the USE of language. More importantly, he posits a principle of the AUTONOMY OF MEANING (ibid: 274-5), which states that the literal linguistic meaning of a linguistic expression does not determine what ulterior purpose it is used for and vice versa. In response to Davidson's declaration that "there is no such thing as a language", Hacking (1986: 447) asks where that leaves the Tarski-style theories of truth which Davidson has long taken to account for linguistic meaning (e.g. in Davidson 1973). But even in "Derangement", Davidson indicates that passing as well as prior theories have finite bases, are recursive, and may be modelled on a Tarski truth definition (1986: 437-8). His point concerning the autonomy of meaning is that these

> ... criteria for deciding what an utterance literally means, given by a theory of truth or meaning for the speaker, do not decide whether he has accomplished his ulterior purpose... (1981: 264)

According to Davidson, this is mainly because it is not the meaning of the linguistic expression, but the person who utters it, that decides what the ulterior purpose for uttering the linguistic expression is:

> ... what is put into literal meaning... becomes available for any ulterior (non-linguistic) purpose -- and even any illocutionary performance. (ibid: 275)[9]
>
> Once a sentence is understood, an utterance of it may be used to serve almost any extra-linguistic purpose. (1975: 164)[10]

Davidson also argues that none of these extra-linguistic ulterior purposes will give the literal meaning of the linguistic expression used. They may help the interpreter *recover* the literal meaning of the linguistic expression used:

> ... it often happens that we can descry the literal meaning of a word or phrase by first appreciating what the speaker is getting at. (1986: 435; see also Wilson and Sperber 1981, and Sperber and Wilson 1986a/95 for the relevance theory account of disambiguation)

But linguistic meaning, being autonomous, cannot be derived from, or defined on the basis of, a particular "standard" use --

> Since the literal meaning operates as well when the use is absent as when it is present, no convention that operates only in 'standard' situations can give the literal meaning. (Davidson 1981: 275)

-- or any of the "endless uses to which a sentence, with meaning unchanged, can be put" (ibid: 271).

The question is: does "Derangement" represent a rejection of the principle of the autonomy of meaning; and if not, then how are the paper's supposedly Humpty-Dumptyesque proposals to be reconciled with Davidson's earlier work? It may seem obvious why there is no mention of the autonomy of meaning in a paper concerned with

---

[9] Davidson adds that an utterance of a sentence with a grammatical mood indicator may "label" itself as a certain illocutionary act, but it does not follow that the utterance *is* that illocutionary act (1981: 275).

[10] Page numbers for Davidson 1975 are from the 1984 reprint.

the apparent phenomenon of "intended meaning" taking over "standard meaning" (1986: 434). But the fact is that, in "Derangement", intended meaning does not actually *take over* standard meaning, in the sense of replacing it in the (prior) theory to which it belongs. Instead, the intended meaning is conceived of as belonging to a separate, distinct (passing) theory brought in by the speaker and/ or hearer. So, in the ordinary case, the intended meaning remains the language user's intended meaning, which the standard meaning of the linguistic expression used neither determines nor is derived from. But in the case of malapropisms etc, the intended meaning becomes a standard meaning in another (passing) theory of linguistic meaning, while the standard meaning remains a part of the current (prior) theory. Both intended and standard meanings, presumably, are autonomous in the way that Davidson in 1975 and 1981 argues that literal linguistic meanings should be.

Furthermore, although it has a "vanishingly small" field of application (Davidson 1986: 442), the passing theory is nevertheless a complete theory of interpretation, i.e. what Davidson initially regarded as a language. Dummett disagrees with this view of passing theories as "massively reduplicating" prior theories (Dummett 1986: 466), arguing that a passing "theory" is nothing more than "a collection of disconnected propositions" (ibid). Davidson, however, insists that the seemingly small, usually temporary adjustment that the language user has to make is not a contribution to what is implicated, but has to do with what linguistic expression a meaning is the meaning of, and therefore constitutes the difference between one theory of interpretation and another:

> ... when a word or phrase temporarily or locally takes over the role of some other word or phrase (as treated in a prior theory, perhaps), the entire burden of that role, with all its implications for logical relations to other words, phrases, and sentences, must be carried along by the passing theory. Someone who grasps the fact that Mrs Malaprop means 'epithet' when she says 'epitaph' must give 'epithet' all the powers 'epitaph' has for many other people. Only a full recursive theory can do justice to these powers. These remarks do not depend on supposing that Mrs Malaprop will always make this 'mistake'; once is enough to

> summon up a passing theory assigning a new role to 'epitaph'. (1986: 443)

In view of the principle of the autonomy of meaning and its bearing on "Derangement", it is also necessary to take a closer look at Davidson's depiction of "words, phrases, and sentences" as having "*new roles*" in the quote given above. In the context of "Derangement" there is something paradoxical about the idea of "new roles" and "new uses". On the one hand it implies that Mrs Malaprop's uttering of 'epitaph' to mean *epithet* is simply another addition to the "endless [non-linguistic] uses" to which 'epitaph' meaning *epitaph* may be put. On the other hand, Davidson makes it quite clear that 'epitaph' for Mrs Malaprop and her interpreter actually belongs to a theory which states that it is a word that literally means *epithet* -- this supposedly being the result of Mrs Malaprop's intention to use 'epitaph' to literally mean *epithet*. The notion of a linguistic expression literally meaning whatever the language user intends or uses it to mean is precisely what Chomsky, Carston and Dummett object to. And yet, although the essence of this view of language is that the linguistic expression's semantic properties are constituted by, or arise from, the use of that linguistic expression, it would seem that Davidson does not entirely avoid the implication that what Mrs Malaprop uses when she utters 'epitaph' and what, say, Davidson himself uses when he utters 'epitaph' are in some way the same entity -- Mrs Malaprop's *new use*, and the new literal meaning *epithet*, notwithstanding.

That Davidson in fact courts this implication by distinguishing Mrs Malaprop's *new use* of 'epitaph' or Donnellan's *new use* of 'glory' (which involves "merely substitution" (ibid: 441)) from the phenomenon of *new words* ("say in Joyce or Lewis Carroll" (ibid)) is symptomatic of the tendency to cling to *phonetic form* as the individuating feature of a linguistic expression, even when its *meaning* has been "substituted" or is simply a matter for doubt. In 4.1 I argued that this tendency is evident in Carston's account of "misuse". In

"Derangement", however, Davidson ought to have avoided it, as it directly contradicts his attempt (via his positing of passing as well as prior theories) to give linguistic meaning the central role in his account of malapropisms. For Davidson, Mrs Malaprop's uttering of 'epitaph' constitutes the use, not of the expression Davidson uses when he utters 'epitaph', or some purely formal Humpty-Dumptyesque construction which just happens to have previously been used to mean *epitaph*, but of an expression which literally means *epithet* in accordance with a proper, albeit passing, theory. In other words, in uttering 'epitaph' Mrs Malaprop uses a linguistic expression whose literal linguistic meaning she actually intends to express, and consequently *uses* together with the other properties of the linguistic expression. Thus "Derangement" could be taken as a response to accounts of malapropisms like Carston's in 4.1, in which the literal linguistic meaning appears to play no part whatsoever in the language user's "misuse" of a linguistic expression.

This reading of "Derangement" is consistent with Davidson 1981, in which he argues that "an action counts as linguistic only if literal meaning is relevant" (1981: 272), indicating that the answer to the question of

> ... whether an activity is interestingly considered linguistic when meanings are not intended to be put to use. (ibid: 273)

is *no*. In fact he goes beyond the idea that language use must involve the language user's cognitive processing and intentional expressing of literal linguistic meaning, making the stronger claim that if one utters linguistic expression L, literal meaning M, with no intention of expressing M, then to all intents and purposes L does not literally mean M:

> ... someone cannot utter the sentence 'Eat the eggplant' with the words literally meaning that someone is to eat his eggplant unless he intends the sentence to have that meaning, and intends his audience to interpret it as having that meaning; but if it is uttered with the intention of uttering a sentence with that meaning, and it does not in fact have that meaning, then it has no linguistic meaning at all. Literal meaning and

> intended literal meaning must coincide if there is to be a literal meaning. (ibid: 271-2)

In "Derangement" he goes even further, arguing that if one utters L with the intention of, and expectation of being interpreted as, literally expressing N, then L does not mean M but N, as specified by another (passing) theory -- L with literal meaning M and L with literal meaning N being two distinct linguistic expressions, say L_1 and L_2, belonging to two distinct theories.

Hence Davidson never talks about the *misuse* of language, in Carston's sense of incorrectly using the phonological/ phonetic form of L_1 to express N as its literal meaning -- firstly, because this does not involve L_1's literal meaning M, and for Davidson is therefore not an instance of the use of L_1; and secondly, because he appears to regard the intention concerning N as necessary *and* sufficient for L_2 to be used. In other words, the issue of misuse never arises in "Derangement", for if the intention to express N as a literal meaning is sufficient for L_2 to be used, then obviously one never fails to use L_2 "correctly" since one always intends to express N.

On the whole I agree with Davidson's view that it is what the language user does with the *linguistic meaning*, rather than the *phonetic vehicle* of that meaning, that counts as language use. But the fact remains that the relation between linguistic semantics and the non-linguistic cognitive entities (thoughts, assumptions, intentions) involved in language use is such that, if one gives meaning precedence over phonetic form, one immediately appears to lay oneself open to charges of Humpty-Dumptyism. Whereas phonetic form is assumed to be linguistic (and only linguistic), linguistic meaning tends to be regarded as having been "[drawn] from the resources of other... faculties of the mind" (Chomsky 1992a: 115-6), or as parts of the "real", truth-theoretic semantics of the central representational system which linguistic expressions possess "by inheritance" (Carston 1998: 64-5). Similarly, any propositional meaning expressed which falls within the

language user's intentions belongs first and foremost to the realm of the Davidsonian "ulterior purpose", as determined by the central cognitive systems, even though it may also happen to be the meaning of the linguistic expression used. Thus the linguistic content is distinguished from the non-linguistic content of the ulterior purpose solely on the grounds of being conventionally associated with a *phonetic vehicle* that is recognisably *linguistic*. It follows from this that, if the phonetic vehicle is secondary to the meaning in the use of the linguistic expression, then there would be no way of telling where linguistic meaning ends and non-linguistic ulterior purpose begins. In fact there would be no way of telling if the non-linguistic ulterior purpose is all there is.

That, of course, is precisely Davidson's point in "Derangement". He does not deny that there is some sort of relation between meaning and vehicle, and that it has the property of what he calls autonomy. But literalness for Davidson is *not* an expression meaning what it means -- i.e. it is not autonomy -- but an expression intended to be interpreted as meaning what it means. And it is because of its autonomy -- which arises, somewhat paradoxically, from its subservient role as input to the central systems which actually effect its use -- that a particular vehicle-meaning relation remains within the bounds of the theory to which it belongs, and has no bearing on the question of what is used or what is literal. Nor does it have any bearing on the possibility that there is some other, equally autonomous, system of vehicle-meaning relations which may also be used literally.

This may be what Davidson means when he remarks that "we must try to pry apart what is literal in language from what is conventional or established" (1986: 434) -- his argument being that what is literal has to do with the particular set of conventions which the language user chooses to use, and not with any one definitive set of rules or conventions. Thus he ultimately rejects Lewis' R, on the grounds that its most important feature is *regularity over time*, suggesting instead that

> It could even happen that every speaker from the start had his own quite unique way of speaking. (1981: 276-7)

This would be equivalent to Carston's remark about "disparate encodings" (1998: 82, fn 2), but for the fact that Carston's "disparate encodings" are *linguistic* while Davidson concludes that his disparate theories of interpretation are *not languages*. I stress once again that Davidson is not denying that language exists, only that the linguistic is a matter of public and antecedent convention. For a fuller picture of his theory of language it is necessary to refer to the conclusion of his 1981 paper (1981: 278-80). Here, he concedes that there is an

> ... element of the conventional, or of the conditioning process that makes speakers rough linguistic facsimilies of their friends and parents... (ibid: 278)

But this, he argues, is purely SOCIAL -- a "practical" matter of "how society bends linguistic habits to a public norm" which tells us nothing about the "linguistic habits" themselves (ibid). It follows from this that any "standard" or "conventional" system of vehicle-meaning relations is the result of social conditioning rather than some genuinely linguistic imperative.

All this is particularly relevant in respect of Dummett's view of language as a social practice. For Dummett, a language is a set of social conventions which link particular linguistic expressions with particular meanings, and which language users have a "responsibility" to (1986: 473). But it is precisely because Dummett's "language" is *social* in nature, that there is nothing to prevent the language user from relinquishing the responsibility and deviating from the conventions -- or, more importantly, from changing the conventions, or setting up new ones. Perhaps Dummett undermines his own supposedly anti-Humpty-Dumpty argument when he maintains that linguistic expressions have meanings "independently of any particular speakers" but, being part of a social

practice, *not* "independently of anything human beings do" (ibid). The distinction between "any particular speaker" and "anything human beings do" is not absolute, but a matter of degree. A few speakers' "imperfect grasp" of a set of conventions could very well lead to a different set of conventions for an entire "community" of language users.

This is especially true if those few speakers happen to be in positions of -- in Carston's words (1998: 82, fn 2) -- "power and authority":

> Of course, the matter of what gets called an error or a misuse is sometimes more a matter of power and authority than anything to do with language itself; see, for example, Chomsky (1987, 29-30)'s discussion of 'livid' which most of us take to encode 'red' or 'flushed' while what it "really means" (i.e. what the dictionary tells us) is 'pale' or 'greyish'. In such a case most of us will (mis)use the word very satisfactorily without anyone having to make any interpretive adjustment. (ibid)

What Carston is saying here is that 'livid' may linguistically encode *flushed* OR *pale*, but whether the use of the expression to express *flushed* or the use of it to express *pale* is the *misuse* depends on whether the "power and authority" resides with the people who take 'livid' to encode *flushed* or those who take it to encode *pale*. Oddly -- or perhaps not oddly at all -- Carston regards the *pale* minority as the people with the "power and authority", and the others, though constituting the majority, as misusing 'livid' en masse. But the main problem, and contradiction, is that 'livid' ENCODES *flushed* for this "powerless" majority -- as Carston's phrase "disparate encodings" implies (ibid) -- and is therefore not misused, but used perfectly correctly by them to express *flushed*. Carston also uses the phrase "*take to* encode" (my italics), which in its turn rather alarmingly implies that the language user, like Humpty Dumpty, may get to decide what particular linguistic expressions encode.

In the light of his discussion of malapropisms and convention in Davidson 1981 and "Derangement", what Davidson refers to as "the conventions of language" and "our

standard method of interpretation" (1981: 279) are conventional or standard, not so much in Lewis' sense of regularity over time, as in the sense of determining certain meaning-form relations which a reasonably large number of people happen to agree about as a result of having been socially conditioned to conform thus. These "conventions", however, are no more than a "starting point" or a "crutch" (ibid) -- the most that social conditioning ensures is that

> ... we may, up to a point, *assume* that the same method of interpretation that we use for others, or that we *assume* others use for us, will work for a new speaker. (ibid, my italics)

There is, of course, no real guarantee that the prior or even passing theories of any two language users will be identical.

At the end of Davidson 1981, Davidson speculates that it is not convention -- or, to be more accurate, the social tendency toward convention -- that is a condition or feature of language, but the other way around:

> I suggest... that philosophers who make convention a necessary element in language have the matter backwards. The truth is rather that language is a condition for having conventions. (1981: 280)

To me this is more interesting and specific than the conclusion of the later "Derangement" (i.e. that convention either constitutes some crucial feature of language that has nothing to do with the vehicle-meaning relation, or is not involved in language at all). Davidson's line of argument is this: in order to share a convention with someone else one has to be able to attribute beliefs and desires to him; and in order to have, and attribute, beliefs and desires one has to have a language. However, if vehicle-meaning relations are merely social and not linguistic, then what and where are the linguistic expressions which are necessary for

beliefs and desires, and how exactly are they related to the social conventions which they facilitate? Davidson raises the question, but does not answer it.

5 Linguistic semantics without "linguistic" conventions: Burton-Roberts' representational conjecture

Davidson concludes that the linguistic does not consist in theories or systems of conventions which govern the relation between literal meaning and that which it is the literal meaning of. Such conventions, he argues (1981: 278-80), are merely the result of social conditioning, and "[do] not explain what is basic to linguistic communication" (ibid: 280). Burton-Roberts (e.g. 1994, 1999a, 1999b) arrives at the same conclusion from the opposite direction, starting from the Chomskian view that the linguistic has nothing to do with *communication*, being a natural and entirely internal state of mind/ brain. However, his is a far more radically internalist theory of the linguistic than Chomsky's, as their respective approaches to phonology clearly illustrate. In chapter 2 I argued that the *I* nature of Chomsky's *I*-language is deeply compromised, chiefly by its interfacing with, and being in part determined by, non-linguistic *performance systems*. With regard to the phonology of I-language, Chomsky claims that

> The special properties of the phonological component relate to the need to produce instructions for the sensorimotor systems, for production and perception [of E-physical utterance phenomena]. (1995b: 229)

He concedes that this may be

> ... the source of other imperfections of [the computational system of I-language], and in this sense "extraneous" to language... (ibid)

BR, on the other hand, insists that the linguistic has no such direct contact with systems that facilitate the production/ perception of E-physical utterance phenomena, and, by extension, with the utterance phenomena themselves. He also maintains that the indirect contact the linguistic does have with performance systems and utterance phenomena is of quite a different nature. He makes the conjecture that E-physical utterances are in what he calls an M-REPRESENTATIONAL relation to I-linguistic expressions. This "representational conjecture" provides an account of the I-E *relation* while maintaining an absolute I-E *distinction*. Such a distinction is essential to the conjecture's presupposition that the linguistic is radically internal and austere. It is also essential to the conjecture's refutation of the (somewhat confused and basically externalist) idea that a linguistic expression is an abstract and autonomous entity which, in virtue of being used, nevertheless becomes or is converted to some physical version of itself, complete with the real (propositional, entertainable) semantics in addition to all the original linguistic properties.

BR's representational conjecture is outlined in section 5.1. The conjecture has important implications for the nature and status of phonology. As Chomsky's account of the phonological component demonstrates, the phonological tends to be regarded as I-linguistic, but also as being determined by non-linguistic elements to do with the production and perception of phonetic phenomena. In 5.2 I discuss BR's argument that the phonological is too closely and inextricably associated with E-physical, non-linguistic phonetics to be I-linguistic itself. This, we shall see, leads back to Davidson's rejection of convention as the basis of the linguistic.

The traditional, virtually consensus view is that the phonological in effect constitutes the conventional vehicle of linguistic meaning. In other words, its relation with meaning is what Davidson's (passing/ prior) theories or systems of conventions govern. But if the phonological and its relation to linguistic semantics/ syntax is, as BR argues, not in fact linguistic -- just as the conventions governing vehicle-meaning relations are regarded by

Davidson to be social rather than linguistic -- then BR's representational conjecture and the radically internalist notion of the linguistic which it is intended to support may provide the answer to Davidson's questions about what and where the linguistic really is. 5.2.1 and 5.2.2 deal with this issue and, more importantly, with its implications for linguistic semantics and the notions of *literal meaning* and *language use*. 5.3 is concerned with its implications for the relation between the linguistic and what Fodor calls the *language of thought*.

5.1 The representational conjecture

Burton-Roberts' 1994 representational conjecture (henceforth RC) is that the relation between linguistic expressions and external utterance phenomena is logically of the same kind as that between, say, a table and a painting of it. On this view, E-physical utterance phenomena are *representations* of I-linguistic expressions, produced by speakers in aid of E-physically representing I-linguistic expressions generated by their mentally-constituted grammars.

BR refers to this notion of representation as *M-representation*. "M" stands for Magritte, and is an allusion to his *La Trahison des Images*, with its slogan *Ceci n'est pas une pipe* reminding us that the painting contains, not a pipe, but a *representation* of a pipe. "M" is also to distinguish M-representation from Chomsky's (1995a, 1995b) "linguistic levels of *representation*", or his (1995b) "PF and LF *representations*" (i.e. π and λ). Chomsky himself explicitly states (1995a: 53) that his use of "representation" is "not to be understood relationally, as "representation of..."" (see footnote 4, chapter 2). BR calls this *C-representation* -- for "Chomsky/ constitutive", since such a representation is constitutive of what it represents. For instance, a "phonological (C-)representation" of a linguistic expression simply *is* the phonology of the linguistic expression. An M-

representation *x* of *y*, on the other hand, is in no way constitutive of what it is a representation of, i.e. *y*. Nor does *x* constitute "an *example*, or *actualisation*, or *realisation*, or *instance*, or *manifestation*, or *exponent* of *y*" (BR and Carr 1999: 393). More particularly, an M-representation *x* of *y* is not a *token* of *y*.[1]

The RC makes a corresponding and equally uncompromising distinction between the production of utterances and the generating of linguistic expressions; between the rules that govern the former and the principles that determine the latter (BR 1994: 195, BR and Carr 1999: 394). As utterances produced by speakers are not linguistic but merely M-representational of the linguistic, so the external behaviour productive of those M-representations is guided by M-representational, not linguistic, rules. The truly linguistic principles, on the other hand, do not "guide" or "licence" anything. They do not "licence" external utterance behaviour. Nor do they "licence" the "production" of linguistic expressions, since the grammar generates linguistic expressions only in the internalist sense of specifying what they are.[2] So to talk in terms of "producing linguistic expressions" or "linguistic behaviour" is, on the representational view, to see the linguistic where it is not -- only E-physical M-representations of linguistic expressions are *produced*, and produced through *external behaviour* involving skills, abilities, habits and (followable) rules of M-representation, not of language.

Furthermore, if the E-physical is M-representational of the linguistic, then the relation between utterances and linguistic expressions must be ASYMMETRIC and NON-NATURAL. M-representation is *asymmetric* in that the representans is logically orientated on the representatum, and not vice versa. According to the RC, E-physical

---

[1] See Bromberger 1989 on "types and tokens in linguistics", and BR 1994 (: 186-91) and BR and Carr 1999 (: 389-91) for problems with the type-token account of the relation between linguistic expressions and E-physical utterance phenomena.

[2] And, as BR (1994: 195) and BR and Carr (1999 : 394) maintain, speakers do not "produce new linguistic expressions" -- though they may produce *M-representations* of linguistic expressions which, like all linguistic expressions, are defined by the grammar, but have not previously been M-represented.

utterances are produced by humans in aid of M-representing what is radically internal and austere, i.e. the linguistic. Linguistic expressions, on the other hand, are not generated by the grammar to be objects of M-representation, or to be "implemented" in E-physical behaviour (BR 1994: 196) -- or, for that matter, to be anything other than their linguistic selves. As BR puts it (ibid), the RC reverses the notion that the linguistic is a means to a behavioural end, suggesting instead that it is the behavioural and E-physical which is a representational means to a linguistic end. This "linguistic end" is

> ... 'in there' as a state of mind/ brain, by its nature innocent of the fact that, 'out there', is behaviour representationally aimed at it as at a target. (ibid)

That there is such behaviour aimed M-representationally at the linguistic is, as BR and Carr claim (1999: 394), not a fact about the nature of the linguistic. Nor is it a fact about the nature of the M-representational utterance phenomena which the behaviour is in aid of producing. M-representation is a *non-natural* relation, and the M-representation of *y* by *x* on a particular occasion depends on (i) the intentions of *x*'s producer that *x* M-represents *y* on that occasion; and on (ii) the assumptions of *x*'s perceiver about those intentions (ibid). In BR 1994 (: 199), the M-representational character is described as being "not immanent in" utterances, but " 'read into' " the physical phenomena by

> ... (rational, relevance-seeking) reader-hearers on the basis of their knowledge of what they take to be the intended *representata*, the content of this knowledge being the mentally constituted constraints governing what constitutes an expression in the language.

Another point about the non-natural M-representational relation between *x* and *y* is that *x* does not have to have any property in common with *y*. M-representation can be ICONIC -- as in the case of Magritte's painting of the pipe, which shares some perceptual (mainly visual) properties with its object, the pipe. But, being non-natural, M-

representation can also be CONVENTIONAL. I am using "conventional" here to mean *RULE-GOVERNED* as opposed to the more specific *AGREED UPON.* (However, see sections 6 and 6.1, where I distinguish between *rule-governed* M-representation and M-representation which is both *rule-governed* and *agreed upon.*) BR observes that M-representation *must* be wholly conventional when there is no overlap of the perceptual domains to which *x* and *y* belong (e.g. the visual M-representation of musical sounds), and when *y* is cognitive rather than perceptual (e.g. the representatum of "5", "527", "%" or "π "). As conceived within the representational framework, the linguistic is radically internal in virtue of being entirely cognitive, and can have no properties in common with the external E-physical utterance phenomena which thereby *conventionally* M-represent it.

It follows from the conventional nature of this M-representational relation that there can be almost any number or variety of representational systems, both within and across perceptual domains (BR and Carr 1999: 395). And since they are all equally M-representational and therefore equally non-natural, no representational system could have a closer or more "natural" relation to the linguistic -- and certainly no representational system could be "more linguistic" -- than another. In this sense, speech is no more "natural" than, say, writing. As Burton-Roberts and Carr argue (ibid: 396), speech has to be distinguished from a natural capacity and disposition to vocalise. The latter is naturally and inherently productive of certain sounds, but for those sounds to be M-representations (of anything, not just of the linguistic) they have to be intended and/ or assumed to be so by their producers and/ or perceivers. Thus speech, in BR and Carr's words (ibid), involves "the *non-natural* exploitation of that natural capacity and disposition [to vocalise], as a medium of (n[on-]n[atural]) M-representation of the linguistic".

So on the conjecture that the linguistic is M-represented by external utterance phenomena, it is the latter -- and the abilities, habits, etc involved in their production and perception -- which are brought to bear upon the former. This deployment of the E-physical is most obviously for the purpose of providing other non-telepathic humans with

indirect -- because merely M-representational -- but E-physical access to what is I-cognitive, i.e. the linguistic. It has to be stressed that the linguistic is NOT constituted in this M-representational relation of the E-physical to the I-linguistic, or in the non-natural M-representational conventions which govern it. The property of being conventional belongs exclusively to the utterance phenomena's M-representational relation to the linguistic, just as linguistic properties belong to the linguistic and not to utterance phenomena. Utterances are cognitively significant, but only in virtue of being intended/ assumed to be M-representational of the linguistic. In other words, the cognitive significance of utterances is *purely M-representational*, and must not be confused with the cognitive value which the M-represented linguistic expressions possess in virtue of their linguistic (and especially semantic) properties.

As BR points out (1994: 194), there may in fact be a *tendency* to confuse and/ or identify M-representations with their linguistic representata. This tendency arises from the fact that linguistic expressions (as conceived within a truly internalist linguistic theory) are never physically encountered, while "the [E-physical M-representation of a linguistic expression], but not what is represented, is so accessible to the conscious mind" (ibid). Whereas E-physical utterances and their representational significance obviously lie within the realm of human consciousness, it follows from the internalist view that the linguistic -- being an innate state of mind/ brain -- is not something to which we have conscious cognitive access (BR and Carr 1999: 399). BR and Carr claim:

> Perceptually and consciously, then, speakers have only the E-physical phenomena and their significance to go on. (ibid)

Thus E-physical M-representations of linguistic expressions are probably the closest that humans can *consciously* get to the linguistic -- which, I assume, is not very close at all. M-representations are by definition far from transparent with respect to the nature and

properties of their representata. As BR and Carr point out (ibid: 397), the M-representations of linguistic expressions are not perspicuous enough to be something from which a complete account of the nature of the linguistic representata may be obtained.[3] But they are at least as perspicuous as their *compositionality* allows them to be (ibid, BR 1994: 198). Conventional M-representation is not wholly arbitrary, for there are conventional M-representations which are *complex*. Unlike *simple* M-representations -- signs consisting only of physical symbols (e.g. {*n*}: one symbol, one sign; {*d-o-g*}: three symbols, one sign), complex M-representations are signs composed of signs (e.g. {{*a-t*}{*w-o-r-k*}}). Their composition is not arbitrary but governed by a system of representational rules -- i.e. a CONVENTIONAL SYSTEM OF PHYSICAL REPRESENTATION (CSPR) (BR and Carr 1999: 397). The composition of a complex M-representation is "representational of [some but not all] aspects of the composition of the linguistic expression represented" (BR 1994: 198). So, as far as their compositionality is concerned, M-representational utterances are perspicuous to the extent that we may make inferences about *linguistic* compositionality from the system of rules which governs *representational* compositionality.

More obviously, M-representational utterance phenomena must be perspicuous enough to

> ... afford some kind of cognitive access to what is represented and to allow for acquisition of that kind of access. (ibid)

Furthermore, it would appear that an M-representational utterance reveals enough about its linguistic representatum for representans to be mistaken for representatum, and for properties of representatum to be mistakenly attributed to representans. In BR 1994, BR 1999b and BR and Carr 1999 (: 397-8), it is claimed that the notion of linguistic

---

[3] Hence the *theoretical* nature of *theoretical linguistics*.

ambiguity arises from the confusion of linguistic semantics with representational indeterminacy.

Before I discuss the representational account of so-called "linguistic ambiguity", however, let me review and revise the notational conventions I have been using thus far. I have been using double quotation marks to indicate utterances, single quotation marks to indicate linguistic expressions, and italics for the semantics of linguistic expressions. I subscribe to BR's representational conjecture, and believe that E-physical or E-instantiable utterances -- e.g. "bachelor" -- are M-representations of linguistic expressions. Furthermore, bearing in mind the nature of the M-representational relation, I take it that the linguistic expression M-represented by an utterance "bachelor" need not necessarily be the linguistic expression which means *unmarried male*. It just so happens that, according to the particular CSPR which I am using at the moment, "bachelor" *does* M-represent the linguistic expression 'bachelor'. In order to avoid the implication that the utterance "bachelor" is somehow inextricably associated with the linguistic expression 'bachelor', and to distance the I-linguistic from the E-physical and M-representational, I shall use braces to indicate M-representations (e.g. {bachelor}). Note that these braces and their contents constitute citations of M-representations and *not of utterances.*[4] An M-representation in itself is not an utterance. It can, however, be uttered. More importantly, it is the M-representation of a linguistic expression -- and never the linguistic expression itself -- which a so-called "utterance of a linguistic expression" is an utterance of.

Returning to the topic of "ambiguity", BR rejects the notion of a linguistic expression being "ambiguous" in the sense of having more than one meaning, e.g. 'flying planes' meaning *the flying of planes* or *planes which are flying*. He suggests instead that *the flying of planes* constitutes the semantics of a linguistic expression 'flying planes$_1$', and *planes which are flying*, the semantics of another linguistic expression 'flying planes$_2$'.

---

[4] In fact, strictly speaking, the braces and their contents constitute "citation[s] of the *alphabetical* strings conventionally employed to represent linguistic expressions" (BR 1994: 203, my italics).

This rejection of linguistic ambiguity is motivated by the view that, in the individuation of linguistic expressions, the semantic has as significant a role as any other linguistic property (BR 1994: 182-3, BR and Carr 1999: 397-8). Thus the idea that there is a single "ambiguous" linguistic expression meaning *the flying of planes* OR *planes which are flying* arises from the attributing of linguistic, semantic properties to the non-linguistic, non-semantic M-representation {flying planes}. This leads to the single M-representation {flying planes} -- with its M-representational indeterminacy as to which of the two linguistic expressions 'flying planes$_1$' and 'flying planes$_2$' is being represented -- being taken for a single "ambiguous" linguistic expression *'flying planes'.

Another of BR's examples of the confusion/ conflation of representational and linguistic properties has to do with syntax (BR 1994: 199, BR and Carr 1999: 398, BR 1999a). BR argues that linear precedence does not constitute a syntactic property of linguistic expressions, but is in fact the intrinsic spatio-temporal linearity of the E-physical phenomena used to M-represent the linguistic (and of E-physical phenomena in general, M-representational or otherwise). To state the argument in another way, using the linguistic instead of the representational as the starting point: entirely *I*-linguistic entities like Heads and Complements cannot be identified with E-physical objects occuring in E-physical strings, just as the *I*-linguistic H-C relation cannot be regarded as a geometric relation involving E-physical phenomena linearly ordered in respect of one another.

It is not just the obvious (M-representational) relevance of M-representations to their linguistic representata, but also the fact that the latter is actually inaccessible to perception and consciousness, that fuels the inclination to identify the M-representational with the linguistic (BR 1994: 194-5, BR and Carr 1999: 398-9). While the distinction between a painting of a pipe and its equally E-physical representatum is perceptually manifest, the fact that the linguistic is cognitive and therefore never physically encountered means that it is only its M-representations, what we do with those M-representations and how we do what we do with them, that are evident to us. In this E-physical absence of representata,

the very fact of M-representation as such, and of specifically M-representational skills and conventions, tends not to emerge. On the other hand, how the painting of the pipe is produced, and what we do (or can do) with it, is obviously not the same as the manufacture and function of the pipe, and therefore easily perceived as being M-representational. Furthermore, the painting may be of value and interest in itself, regardless of its M-representational nature. In contrast, the utterance is more easily confused with its linguistic representatum because the M-representation in this case is of interest to its producers and perceivers primarily (and indeed only) for its M-representational relation to the linguistic.

5.2 What is linguistic semantics the semantics of?: the nature and status of phonology within the representational framework

The representational conjecture (RC) bears upon the issues of literal and linguistic meaning mainly through its implications for phonetics and phonology. The traditional idea of linguistic expressions as sound-meaning pairs (taking the form of (π, λ) pairs in Chomsky 1995b, for example) appears to have given rise to a certain ambiguity with regard to the question of what linguistic semantics is the semantics of. On the one hand, the linguistic semantics is regarded as the semantics *of the linguistic expression*, in the sense of being the "meaning" part of the sound-meaning pair which the linguistic expression is supposed to be constituted by. On the other hand, and yet by the same token, it is also regarded as the semantics *of the phonetic form*, in the sense of being the "meaning" which the "sound" carries or constitutes the vehicle of. For example, Chomsky (1995b: 2) explicitly refers to "sound with a meaning". In fact this notion of the "meaning" carried by some vehicle appears to be fundamental to the concept of literalness. Linguistic expressions may be described as having literal meanings, but strictly speaking it is the

"sound" (phonetic form) or the vehicle -- rather than the linguistic expression, the entire sound-meaning pair -- which the literal meaning is assigned to (by a particular linguistic system), and literally *the meaning of*.

Of course, according to the RC, *sounds* can in no way belong within the wholly I-cognitive domain of the linguistic. Assuming that it has to do with "the acoustic and the articulatory behaviour usually required to produce it" (Burton-Roberts and Carr 1999: 399)[5] -- i.e. E-physical utterance phenomena and utterance behaviour -- the *phonetic* is, in RC terms, not linguistic but M-representational of the linguistic. Whether *phonology* is of the linguistic or the M-representational is a more complex issue. The position which Burton-Roberts ultimately takes -- in BR and Carr 1999 as well as BR 1998 and 2000 -- is that the phonological is *not* I-linguistic but is, together with the phonetic, *M-representational of the linguistic*. The implications of this position for the concept of literal meaning and the nature of the linguistic will be discussed in the remaining sections of this chapter. This section is mainly concerned with the BR and Carr's 1999 discussion of, and arguments *against*, the notion of an I-linguistic phonology.

It is stated in BR and Carr 1999 (: 399), however, that the RC "does not, of itself, deliver a decision on [questions about the nature and status of phonology]". The common assumption is, of course, that the linguistic does include the phonological. The other common assumption is that the phonological in particular (as opposed to the syntactic or the semantic) is associated with the phonetic. BR and Carr 1999 take these assumptions seriously enough to discuss them in connection with the RC, concluding that *a phonology which is I-linguistic would have to be regarded as CONVENTIONALLY M-represented by the phonetic*. In other words, on the view that the linguistic includes the phonological,

---

[5] Articulatory behaviour is *usually* but *not necessarily* required for the production of the acoustic. As BR and Carr point out (1999: 395), it is not the articulatory behaviour but the acoustic which actually M-represents the linguistic, and "how [acoustic M-representations] are produced (be it by vocal articulations, including those of ventriloquists, or by key-pressing on a speech synthesiser) is irrelevant".

the phonetic would be M-representational of the linguistic in virtue of specifically M-representing its phonological elements.

However, the notion of a phonology which is I-linguistic and conventionally M-represented by the (clearly E-orientated, non-linguistic) phonetic has two problematic implications. Firstly, such a phonology must be entirely lacking in phonetic content. It certainly would not be *phon*ology, in the sense of involving E-physical *phones* (or even abstractions from E-physical phones). Nor would it be *mentalistic phonetics*, or *phonetically interpretable*. As BR and Carr point out, a radically internal and austere phonology that is conventionally M-represented by the phonetic must be

> ... 'abstract' beyond anything dreamed of even in the most 'abstract' phonology. (1999: 400)

Secondly, an I-linguistic phonology would not only be in a conventional relation with the M-representational phonetic, but also with the I-linguistic syntax and semantics. The particular relations between phonological objects, syntactic rules and semantic forms constitute the main individuating feature of a "particular language" (e.g. English or Swahili, following Chomsky 1995a: 13). So the way in which the phonology is aligned with the syntax and semantics must be as arbitrary as is consistent with the diversity of "particular languages". In short, there must be different sets of phonology-syntax-semantics tuples for different "particular languages".

One obvious probem, arising from the first implication, is this. Apart from the fact of the I-linguistic phonology's conventional relations with the other I-linguistic (syntactic/ semantic) elements, only negative inferences -- mainly to do with there being absolutely nothing phonetic about such a phonology -- can be made about its I-linguistic nature and rationale. Certainly there is nothing that can be ascertained about what distinguishes it from the syntax or semantics, and especially about what sets it apart as the specific M-

representatum of the phonetic. Even more problematically, what *can* be ascertained about an I-linguistic phonology -- i.e. the conventional nature of its I-linguistic relation to the syntactic/ semantic as well as its non-linguistic relation to the phonetic -- is precisely what makes it seem less than *I*. The non-naturalness of the linguistic relation to the syntax and semantics is at odds with the notion of the I-linguistic as a natural object, a state of mind/ brain. As for the non-linguistic relation to the phonetic: that the phonological happens to be the one element of the linguistic which the phonetic is assumed to M-represent -- or which, if a "particular language" has no phonetic M-representational medium (e.g. "sign languages"), is *not M-represented at all* -- suggests that it is more closely associated with the E-phonetic than an I-linguistic phonology is supposed to be.

BR and Carr claim that there are other factors, quite apart from RC considerations, which support the argument that the phonetic would have to be conventionally related to the phonological if the latter is linguistic (1999: 400). One has to do with the substantive universals of such a phonology. BR and Carr argue that only the *strong* notion of universal applies in this case. The reason is as follows. In the words of BR and Carr, a *weak* notion of universal is

> ... the idea of a universal bag of properties from which "particular languages" may differentially select. (1999: 401)

and a *strong* notion of universal is

> ... the idea of a highly constrained set of properties, not given in advance of enquiry, actually attested in *all* "particular languages". (ibid)

BR and Carr claim that there cannot be weak *linguistic* universals in the phonological domain, because it is specifically and peculiarly in the phonological domain that the weak universals are obviously *not linguistic at all*. As BR and Carr point out, weak universals

like 'ingressive' and 'pharyngeal' are clearly determined by "human physiology between nose and larynx" (ibid). In short, they are universals "not of language but of human physiology" (ibid). This leaves us with the strong notion of universal, and the exclusion from the phonological of physiologically orientated properties like 'pharyngeal' which are not attested in all "particular languages". The nature of the phonological, if linguistic, must therefore be so absolutely different from that of the phonetic that the relation between the two can only be conventional.

Furthermore, any natural or more fundamental relation between the phonological and the phonetic is ruled out by the fact that, whereas the phonological is -- on the view that phonology is linguistic -- a feature of every "particular language", not all of these "particular languages" have a phonetics. As I mentioned above, what we call "sign languages" are clearly non-phonetic. Thus the so-called "strong universals" in the phonetic domain are simply *not strongly universal* -- the phonetic is just one of the several media of M-representation, and therefore in no position to correspond to or determine "phonologically" strong universals attested in *all* "particular languages".

All these problems and difficulties may be due to the strong possibility that the very idea of phonology -- and of various "particular languages" arising from the permutations and combinations of the phonological with the syntactic and semantic -- actually has an externalist basis, and is therefore completely at odds with the internalist theory in which BR and Carr 1999 attempts to incorporate it. Thus the discussion of the difficulties of locating and defining phonology within the internalist framework in BR and Carr 1999 could be said to pave the way for BR 2000, which makes no attempt to keep the notion of an I-linguistic phonology. Instead, it starts from the position that the notion of phonology is only compatible with what BR calls the *essentialist* or *generic* view that the linguistic is constituted by a variety of "languages" (with a small "l") -- i.e. "particular languages" like English or Swahili. Each of these "languages" has a phonology in order that its "expressions" may be E-physically instantiated for the purpose of

communication. In other words, the phonology of a "particular language" is quite literally a *phon*ology involving E-physical *phones*, and therefore necessary for the "language" to function as an E-physical medium for communication.

On the other hand, there is the *naturalist* view, according to which the linguistic is completely internal in virtue of being a natural, innate state of mind/ brain -- i.e. Language (with a capital "L") rather than "(particular) languages". Unlike "languages", Language is only E-physically instantiated as states of brain in individuals. It is not functional -- not for being used, not *for* anything. BR subscribes to this truly internalist view of the linguistic; and to the notion that phonology, being E-realisable for the particular *function* of carrying/ conveying meaning, cannot be I-linguistic. As BR and Carr point out:

> On this alternative, there is no reason why the phonological should not be (unschizophrenically) *phon*ological -- "mentalistic phonetics", to do with the production/ perception of "sounds as such". (1999: 402)

In other words, this leaves the phonological free to be as closely related to the phonetic as it can get, to be orientated toward the phonetic rather than the other M-representational way around. In short, on these terms phonology is not linguistic but M-representational.

On this view, the phonological is internal only in the sense of being

> ... an internal*ised* attitude to the acoustic phenomena produced by fellow humans in the local community in aid of M-representing the linguistic. (BR and Carr 1999: 402)

That is, a phonology consists in regularities and categories abstracted from what acoustic phenomena are used, and how they are used, by a particular group of people to M-represent the linguistic. BR and Carr 1999(: 402) concludes that a phonology is in fact *a Conventional System of Physical Representation (CSPR) for M-representational media which are phonetic*. BR 1998 is more specific: it states that phonology concerns the

categories abstracted from the M-representational use of the acoustic medium; and morphology, the M-representational conventions defined in terms of those categories.

What is more, BR argues that a "particular language" with a phonetics does not just include, but IS, a morphophonological system -- i.e. a CSPR consisting of conventions abstracted from, and governing, a particular use of a particular set of acoustic phenomena to M-represent the linguistic. This means that a "particular language" with a phonology and phonetics bears exactly the same relation to the linguistic as, say, "sign languages" or systems of writing (even those which are acquired after, and parasitic on, a phonology): all are CSPRs, though involving different E-physical media, used for the M-representation of a unique cognitive system, namely *Language*.

BR's proposal that "particular languages" are CSPRs is also consistent with the notion of "language acquisition". That the linguistic has to be learned or acquired is obviously incompatible with the internalist view of it as natural and innate. For BR and Carr 1999 and BR 1998, however, what is learned/ acquired is not Language in the naturalist sense but "particular languages", i.e. conventional systems of physical representation (CSPRs) for the external M-representation of a unique Language. Since it is an internal*ised* attitude to M-representations, a CSPR or "particular language" can only be acquired (mentally internalised) when the relevant (M-representational) E-physical phenomena are actually encountered or experienced. Even before that, however, the I-linguistic representatum -- i.e. Language -- must be cognitively accessed in some unconscious way, in order for the individual to be able to ascribe representational intentions and thereby pick out those E-physical phenomena which are (or are assumed to be) intended as M-representational of the linguistic.

5.2.1 M-representation and the conventional vehicle-meaning relation

In my opinion, the naturalist view and the representational conjecture (RC) provide some answers to the questions raised in Davidson 1981 and 1986 about how conventions governing sound-meaning (or vehicle-meaning) relations bear upon the linguistic and vice versa. Davidson rejects the view of the linguistic as including systems of such conventions (i.e. prior and passing theories), learned in advance and shared by speakers. His grounds are that speakers are constantly constructing different systems of conventions for different circumstances (e.g. different interlocutors). Thus the main argument of Davidson 1986 is that the term "language" should not be used to distinguish something -- i.e. the setting up and application of supposedly "linguistic" conventions -- which he regards as indistinguishable from "knowing our way around in the world generally" (1986: 446); and, more specifically, from the socially conditioned tendency to "speak much as [one's] neighbours do" (1981: 278). Furthermore, he suggests that it is not the conventional that is a feature of language, but the linguistic that is necessary for the establishing and sharing of conventions.

So Davidson does assume that there is something that the term "language" refers to, but argues in Davidson 1981 and 1986 that, whatever it is, it is independent of -- though necessary for -- the sociocultural phenomena which he considers prior and passing theories to be. This may not seem particularly interesting for someone who subscribes to the internalist view of the linguistic. On this view there is a distinction between the linguistic and the *use* of it in the production/ interpretation of utterances. Prior/ passing theories obviously arise from the latter rather than constitute the former. Thus Chomsky remarks (1992a: 119) that Davidson 1986 "[does] not seem to show very much". For Chomsky there is of course an I-language, a generative procedure, distinct from the interpretation faculty which involves all the capacities of the mind, and is involved in "everything that people are capable of doing" (ibid: 120).

However, the fact remains that the non-natural sound-meaning relations governed by Davidson's systems of conventions are also crucial to, and problematic for, Chomsky's I-language. In the I-linguistic lexicon, lexical items have entries containing information about the "sound-meaning relation" (Chomsky 1995b: 236), upon which the computational component computes to give (π, λ) pairs that determine "the phenomena of sound and meaning for [a particular I-language L]" (ibid: 224). Most importantly, the supposedly I-linguistic (π, λ) is what it is, a sound-meaning pair, in virtue of being specifically accommodated to the A[rticulatory]-P[erceptual] and C[conceptual]-I[ntentional] performance systems which effect the process of *utterance interpretation*.

Furthermore, although it is supposed to be the basis for the construction of prior/ passing theories during utterance production/ interpretation, Chomsky's I-language is similar to those very theories in two ways. Firstly, Chomsky and Davidson seem to agree that systems of sound-meaning relations -- prior theories for Davidson, I-linguistic for Chomsky -- differ from speaker to speaker. Davidson suggests:

> It could even happen that every speaker from the start had his own quite unique way of speaking. Something approaching this is in fact the case, of course. Different speakers have different stocks of proper names, different vocabularies, and attach somewhat different meanings to words. (1981: 276-7)

The "unique way of speaking" and "different stocks of proper names" may be regarded by Chomskians as obviously having to do with so-called "*language use*", but "different vocabularies" and attaching "different meanings to words" is consistent with the fact that -- the "I" in "I-language" being also for "individual" -- a "particular I-language" for Chomsky is the I-language of a particular person. Therefore a "particular I-language" is not the same as the "particular languages" -- e.g. English, Swahili -- of BR and Carr 1999. Chomsky claims that

> In the empirical study of language, it has long been taken for granted that there is nothing in the world selected by such terms as "Chinese", or "German", or even much narrower terms. Speaking the same language is much like "living near" or "looking like"... (1995a: 48-9)

English, Chinese, etc are "community norms" or "social practices", not I-languages. Nor is any other "standard language" or system of terms invented by authority figures or experts like dictionary writers and scientists. A particular I-language is what Chomsky says *he* has -- one that is different from the particular I-language of Boris, "a monolingual speaker of some variety of Russian" (1992a: 123); and even from the particular I-language of Jones, although "I can understand Jones, within limits, because my I-language is not too different from his" (ibid).[6]

The second feature common to I-languages and passing/ prior theories has to do with Chomsky's claim that the I-language is a "steady state" of the language faculty:

> ... the cognitive system of Jones's language faculty *is* modified in response to linguistic experience, changing state until it pretty much stabilizes, perhaps as early as six to eight years old, which would mean that later (non-lexical) changes that have been found, up to about puberty, are inner directed. (1995a: 13)

Whereas "later (non-lexical) changes" are described here as "inner-directed", the "steady state" (i.e. I-language) does undergo later *lexical* changes, and does so in response to what Chomsky calls "linguistic experience". For instance, the modifications made by an individual to his *usage* of particular expressions to coincide with the *usage* of other individuals is described by Chomsky as "marginal changes of *I-language*" (1992b: 217, my italics) and "marginal and rather arbitrary variations of *I-language*" (ibid: 228, my italics again). No doubt these are "marginal" and very small changes; but so are those

---

[6] For problems with admitting "individual languages" but not English, Swahili, etc as I-linguistic, see BR and Carr 1999(: 388-9).

which, for Davidson, amount to the replacing of prior theories with passing theories (or other prior theories).

From his 1998 and 2000 position on the nature and status of phonology, Burton-Roberts' views on "lexical change" and the lexicon (as usually -- morphophonologically -- understood) may easily be deduced. The representational conjecture (RC) situates E-instantiable phonology, as well as the conventionality of its association with linguistic semantics (i.e. "the sound-meaning relation"), in the domain of the M-representational. So according to the RC and the naturalist view, the linguistic

> ... cannot be thought of as including "lexical items" as generally (i.e. morphophonologically) understood. On the contrary, so understood, the lexicon precisely is the locus of extralinguistic, M-representational (local, culture-bound) arbitrariness. It has to be acquired and stored. (Burton-Roberts and Carr 1999: 404)[7]

And with regard to Chomsky's and Davidson's problem: by excluding the conventional/ arbitrary from the linguistic, the RC makes it possible for the linguistic to be regarded as a completely internal and natural state of mind/ brain, "universal" in the sense of ranging over all the individuals innately endowed with such a mind/ brain state, and entirely

---

[7] Note that in 1994 BR had yet to conclude that phonology and the (morphophonological) lexicon are M-representational rather than linguistic.Thus he assumed that *words* were the basic units from which *sentences* and *phrases* were composed, and that the complexity of M-representations was M-representational of this lexical compositionality:

> ... the usual convention -- common to orthographic and acoustic representation -- is that the physical representation of a complex expression is composed of ordered physical representations of the words that the sentence or phrase contains. As a representation of a sentence (or phrase), then, an utterance is representational of at least (and perhaps at most) the lexical properties of the linguistic expression it is a representation of. (1994: 198)

But it follows from the exclusion of morphophonology from the linguistic that words, and the complex structures constructed from words, cannot be linguistic. Besides, it is clear that what constitutes a word varies from particular language to particular language, i.e. from CSPR to CSPR; and that the syntactico-semantic object to which a particular word (M-representationally) corresponds may be conceptually complex, and far from being a basic linguistic unit.

independent of the external behaviour and sociocultural circumstances which necessitate the construction and implementation of so-called "linguistic" (actually M-representational) conventions, and which also cause variations in, and changes to, those conventions.

More specifically, the RC's removal of phonology and (morphophonological) lexicon from the linguistic means that linguistic expressions -- i.e. what is generated by the innate, unique grammar which constitutes the linguistic -- are in fact *exclusively syntactico-semantic* I-linguistic objects (BR and Carr 1999: 403-4).[8] Note that *syntax* in this naturalist sense excludes many properties that are traditionally regarded as syntactic -- namely those which are associated with the E-physically instantiable morphophonology. These properties are consigned by the RC to the domain of M-representation.

This shift of the phonological and its associated morphosyntactic features from the linguistic to the M-representational may be described in terms of Chomsky's Minimalist Program, perhaps as a way of converting Chomsky's "I-language" -- which, as it is, seems very much like a CSPR -- into something that is consistent with the internalist standards of BR's (and, ideally, Chomsky's own) naturalist view. A grammar without a phonology is one without a phonological component. And without the phonological part of C_{HL} there would also be no Spell-Out: the operation which applies at the point of the computation where the overt component splits into the phonological and covert components, and which strips away from the structure formed by the overt component the elements relevant to the PF representation and the phonological component. It follows from this that morphosyntactic properties -- i.e. overt PF syntax -- would share the non-linguistic,

[8] I have argued that the lexicon, as the locus of the Saussurean arbitrariness of vehicle-meaning pairs, is in fact an M-representational phenomenon. However, it is necessary to distinguish the lexicon of a particular CSPR -- or a *vocabulary*, as BR (2000) calls it -- from the unique *Lexicon* of the I-cognitive, vehicle-less linguistic. The Lexical items of the Lexicon are vehicle-less syntactico-semantic objects generated by the grammar. Note that the characteristics of lexicons/ vocabularies must not be projected onto the Lexicon. For a start, concepts of what is simple and what is complex, as applied to *E-instantiable* words and sentences, clearly cannot be brought to bear upon *I-cognitive* syntactico-semantic objects. BR also points out (ibid) that there is no reason to expect the (M-representational) relation between vocabulary items and Lexical items to be necessarily isomorphic.

M-representational status of phonology, morphology and the (morphophonologiical) lexicon.

Exactly what these properties are is the subject of ongoing research, although there is some relevant material in BR 1999a (on linear precedence, mentioned in 5.1) and BR 1998 (which argues that case and agreement are not interpretable at LF precisely because they are not linguistic but M-representational). What I am concerned with here, however, are the syntactico-semantic objects which the RC reduces linguistic expressions to, their bearing on the notion of linguistic semantics, and associated issues involving the notions of "literal meaning" and "language use".

If syntax consists only of covert LF syntax, i.e. syntax serving the linguistic semantics -- or what Davidson (1981: 279) describes as "the pattern of inference and structure created by the logical constants" -- then the linguistic expressions generated by BR's innate, unique grammar would in effect be Chomsky's λs untrammelled (as it were) by πs. Alternatively, they could be regarded as logical forms which -- contrary to Sperber and Wilson's relevance theory -- are *never encoded*, because there would be no place for the notion of "*linguistic* encoding". It follows from this that linguistic semantics is, strictly speaking, not *meaning* in the sense of *being the meaning M of some vehicle U.* Or rather, it is not meaning in any relational sense of "meaning".[9] Furthermore, linguistic semantics is not the basis of LITERAL MEANING (whether of "linguistic expressions" or "utterances"), since the very idea of literalness has to do, not with M alone, but with M being -- and being taken to be -- the meaning of what it is literally the meaning of, i.e. U. What the RC highlights is the fact that the U-M relation is the locus of non-natural

---

[9] Nevertheless, I will have to continue on the lines of linguistic expressions *having meanings*, or *meaning whatever they mean*, until the controversial issue of *whether linguistic expressions in fact constitute meaning (in a non-relational sense)* is addressed in section 5.3.

Saussurean arbitrariness and convention. No natural object is inherently a Saussaurean signifier. Objects are linked with concepts or ideas in Saussurean signs if and only if they are *intended, or assumed to be intended*, to "signify" those concepts or ideas.

So a more precise way of putting Davidson's rather confused 1981 claim --

> Literal meaning and intended literal meaning must coincide if there is to be a literal meaning. (1981: 271-2)

-- is to say that "intended literal meaning" is the only sort of literal meaning there is. Much closer to this non-natural notion of literal meaning is Davidson's 1986 argument that the intention to express M with "expression" U is necessary and sufficient for *"expression" U with literal meaning M* to be used -- bearing in mind Davidson's conclusion that the passing and prior theories which determine U-M/ vehicle-meaning relations *do not constitute the linguistic*. The RC manages to explain what these non-linguistic, non-natural vehicle-meaning relations involve, and especially how they bear upon linguistic expressions generated by a *natural*, mentally constituted grammar. According to the RC:

> (1) What is generally assumed to be an *E-physical vehicle U of meaning M* -- or signifier U, or whatever M is literally the meaning of -- is actually *an E-physical object that is produced, or assumed to be produced, with the intention of conventionally M-representing a linguistic expression.*
>
> (2) What is generally assumed to be the LITERAL MEANING *M of U* is actually *the particular cognitive value which the E-physical U possesses, or is assumed to possess, in virtue of M-representing a linguistic expression with a particular semantics.*

Thus it follows from the RC that the concept of literalness and literal meaning -- together with the U-M relation upon which it is based -- actually has to do with non-natural M-representation rather than with the linguistic. For the RC, an E-physical object U cannot strictly be regarded as "a vehicle of meaning M" or as "having meaning M", if only because this relational notion of "meaning" seems to imply that U is naturally and intrinsically possessed of meaning. The fact is that U is *not* naturally and intrinsically possessed of meaning. To think that it was would be to project onto the M-representation that which is only true of what it M-represents.

On the other hand -- as I have already pointed out -- linguistic expressions cannot be regarded as "having meanings" either. Linguistic expressions are naturally and intrinsically possessed of a semantics, in virtue of being at least partly (and possibly wholly) constituted by that semantics. So, even within the RC framework, it does seem reasonable to retain a notion of *literal meaning M of vehicle U*, as long as it is clear that M is NOT the semantics of a linguistic expression U, but *the non-natural M-representational significance of an M-representation U of a linguistic expression.*

It also follows from (1) and (2) that:

> (3) What, for Chomsky or Carston, is the literal USE OF A LINGUISTIC EXPRESSION to express its literal meaning is -- according to the RC -- *the use of a non-natural E-physical M-representation of some linguistic expression.*

More specifically: what tends to be regarded as the literal use of a linguistic expression is (for the RC) the production and manipulation of an *E-physical* M-representation of some linguistic expression, in aid of *E-physically* M-representing that very linguistic expression. This is consistent with my argument in 4.1, that what "linguistic expression" is used is generally taken to be the "linguistic expression" *whose (E-physical) phonetic form is produced* rather than the "linguistic expression" whose meaning is intended to be

expressed. On the RC the phonetic form which is uttered is not of course linguistic but M-representational of the linguistic. Furthermore, what linguistic expression this phonetic form M-represents depends on what *C*onventional *S*ystem of *P*hysical *R*epresentation (CSPR) its utterer is using. Therefore the *literal meaning* (i.e. M-representational significance) of the phonetic form, and the *literal use* of the phonetic form to express its literal meaning, are both determined by *M-representational convention.*

Although I have been using "conventional" in the sense of *rule-governed*, the term also indicates *consensus.* In an ordinary social context, what is generally regarded as constituting the literal meanings of the expressions of a "language" has to be agreed upon, or at least assumed to be agreed upon, by the members of the social group which uses that "language". Thus CSPRs and the "literal meanings" which they govern tend to be shaped by what Davidson calls the socially conditioned tendency to "speak much as [one's] neighbours do" (1981: 278), and other sociocultural factors.

It is from this sociocultural aspect of M-representation that Chomsky's and Carston's concept of "language MISUSE" arises:

> (4) What Chomsky and Carston regard as the MISUSE OF A LINGUISTIC EXPRESSION is, on RC terms, the *misrepresentation of a linguistic expression*, where "misrepresentation" means an M-representation that does not conform to the CSPR agreed upon by the members of the relevant social group.

In other words, what is actually being misused is not a linguistic expression but an M-representation. The important point is that it is only *with respect to a particular CSPR of a particular community* – especially one which is endorsed by authority figures, and recognised (if not actually strictly adhered to) by a large proportion of the community members -- that M-representations of linguistic expressions can be said to be *misused*, or that linguistic expressions can be said to be *misrepresented.* To misuse an M-

representation of a linguistic expression is in effect to deviate from some CSPR -- to use an E-physical object which, according to that CSPR, conventionally M-represents one linguistic expression, to M-represent another linguistic expression.

Actually I am not sure if Mrs Malaprop, for example, could be said to have *deviated from* an M-representational rule which she appparently has no knowledge of, or to have *misused* an M-representation which she has not in fact used. According to the CSPR which she chose to use on that occasion, {derangement} *IS* the M-representation of 'arrangement'. Strictly speaking, that is the M-representation which Mrs Malaprop intended to (and did) use, and the convention (and corresponding CSPR) which states that {derangement} M-represents 'arrangement' is the convention (/ CSPR) which Mrs Malaprop intended to (and did) follow. Similarly, Humpty Dumpty does not deviate from or misuse one CSPR so much as choose to use another. Unlike Mrs Malaprop, however, he is aware of the M-representations and M-representational rules which he is expected to (but deliberately does not) use. So my point is this. A person always uses an M-representation "correctly", i.e. in aid of M-representing what that M-representation M-represents, because what the M-representation M-represents -- *what M-representation it is* -- is determined by the CSPR which he himself chooses to adhere to on that particular occasion.

What I have been arguing so far is that the concepts of *literal meaning* and *use/misuse*, as usually understood, are actually M-representational concepts rather than linguistic concepts. Furthermore, what constitutes the "literal meanings" of M-representations of linguistic expressions, and what constitutes the use or misuse of those M-representations, is determined by M-representational conventions which tend to have a sociocultural basis. Linguistic expressions, on the other hand, are (I argue) naturally and at least partly constituted by their semantics. As purely I-cognitive objects IT IS OBVIOUSLY IMPOSSIBLE FOR LINGUISTIC EXPRESSIONS TO BE *USED* IN THE SAME SENSE AS THEIR E-PHYSICAL M-REPRESENTATIONS ARE *USED*

(if, indeed, linguistic expressions can be said to be *used* at all). And it follows from this naturalist view of the linguistic (Burton-Roberts 1998, BR 2000), and from the RC, that IT IS ALSO IMPOSSIBLE FOR LINGUISTIC EXPRESSIONS TO BE MISUSED. In another context, BR remarks (1994: 184) that a so-called "ambiguous" linguistic expression which a speaker utters is actually not ambiguous at all:

> After all, the *speaker* knows which (necessarily unambiguous) linguistic expression she is uttering.[10]

For the same reason, there is no way a speaker can *misuse* a linguistic expression. She may *misrepresent* a linguistic expression by using what -- according to some Conventional System of Physical Representation (CSPR) -- is the wrong M-representation. But the linguistic expression which she actually M-represents (whatever E-physical object she uses or misuses as the M-representation) is always the linguistic expression which she intends to M-represent, since she cannot but know exactly what linguistic expression she intends to (and *does*) produce an M-representation of. If she did not know what linguistic expression she intended to M-represent, the E-physical phenomena she produces would not be M-representational of any linguistic expression.

The accounts of "misuse" and "ambiguity" set out above would be Humpty-Dumptyesque -- centring as they do on *speaker's intentions* -- were it not for the fact that the putative "expressions" which mean what one intends them to mean are not linguistic expressions but E-physical phenomena used in aid of M-representing linguistic expressions. The RC account of ambiguity may seem less Humpty-Dumptyesque because, whether it is 'flying planes$_1$' or 'flying planes$_2$' which she intends (flying planes) to M-represent, the speaker would be abiding by the convention -- established by a certain

---

[10] Here, "to utter linguistic expressions" is of course elliptical for "to produce E-physical, non-linguistic utterances in aid of M-representing linguistic expressions". See Burton-Roberts 1994: 196-7.

group of people and endorsed by the relevant authority figures -- which in effect states that {flying planes} is the M-representation of both 'flying planes$_1$' and 'flying planes$_2$'.

Though she does not know it, Mrs Malaprop's intentions are directed towards a CSPR which is different from the CSPR being employed by her interlocutor (and by the audience/ reader of the play) -- i.e. the CSPR which states that {derangement} M-represents 'derangement' and not 'arrangement'. This disjunction between two different CSPRs creates the impression that what {derangement} "means" can be bent, Humpty Dumpty fashion, to Mrs Malaprop's will, as a result of Mrs Malaprop's use of {derangement} to "mean" *arrangement*. But this is actually a perfectly ordinary, expected and not very serious occurrence in M-representational terms. According to the RC, what {derangement} *"means"* is merely the M-representational significance of the E-physical object {derangement}, and Mrs Malaprop is simply using a CSPR according to which {derangement} M-represents the linguistic expression whose semantics is constituted by *arrangement*. Besides, Humpty Dumpty himself is guilty of nothing more serious than deliberately using a CSPR which he knows is *not* the CSPR which Alice adheres to.

Actually, there is a sort of circularity about the above argument: namely, that in order for a linguistic expression to be M-represented by some E-physical object, it is necessary and sufficient that the E-physical object is *intended* to be M-representational of that linguistic expression. This is because what one "knows" about the linguistic expressions which one intends to produce M-representations of -- how one knows which linguistic expression is M-represented on a particular occasion -- is in fact a matter of what we deduce from the M-representations and CSPRs themselves. So when I (and BR) say that Mrs Malaprop *knows* what linguistic expression she produces an M-representation of, I am using "knows" in a rather broad sense. As I mentioned in 5.1, any direct cognitive access a person has to the mentally constituted grammar and what it generates[11] is *not*

[11] An innate, mentally constituted grammar is what the linguistic is assumed to consist of. But the linguistic expressions which this grammar generates are *not mentally constituted* -- or at least not

but can justly at least be thought of conscious, right?

conscious. What *is* directly accessible to the consciousness is the E-physical and E-perceptible. E-physical objects are what humans can consciously perceive and identify, as well as mentally represent. We cannot consciously pick out the linguistic expressions generated by the grammar, but we *can* voluntarily manipulate the E-physical M-representations of those linguistic expressions. We can do this physically, and in doing so engage in an activity which is traditionally regarded as the use of "linguistic expressions" (actually M-representations) in external behaviour (especially communication). We can also do this mentally, which is what Dummett (1986: 470-1) describes as the use of "linguistic expressions" as "vehicles of thought". Dummett observes that "language"

> ... is a vehicle of thought because it is an instrument of communication, and not conversely. (ibid)

This idea, for which Dummett (in my view) fails to provide a coherent theoretical basis, is given substance by the representational conjecture. In M-representational terms, what serves as an "instrument" or "vehicle" for Dummett is first and foremost an E-physical M-representation which can be used in E-behaviour as an "instrument of communication", before being mentally internalised as a "vehicle of thought".

Thus BR and Carr (1999: 403) argue that the acquisition of conscious cognitive access to the linguistic and the acquisition of a CSPR are "mutually dependent". And on the assumption that in order to *know* something one has to be in a state of consciousness: one's *knowledge* of the nature of the linguistic expression which one intends to M-represent, as determined by the conscious cognitive access which one has to its properties, is necessarily mediated by the E-physical M-representation of the linguistic expression, and is therefore only as comprehensive and exact as the M-representation (and

---

in the way that the grammar, as an element of the *brain*, is mentally constituted. As BR puts it (1994: 209 fn 17): "They are cognitive only in virtue of being constituted in the generative capacity of a grammar and of my assumption that grammars are mentally constituted. Linguistic expressions are the generatve epiphenomena of a grammar (but not less real for that)."

the CSPR to which it belongs) is perspicuous. In other words, I am assuming that it is in fact very limited and indirect (as I mentioned in 5.1). In contrast, the individual's innate (pre-CSPR) and *unconscious* cognitive access to the linguistic is very possibly complete as well as direct. This sort of cognitive access is necessary for CSPR acquisition, the reason (as I mentioned in 5.2) being that the individual is only able to assume that particular E-physical phenomena are (intended to be) M-representational of the linguistic if he has some sort of prior access to the representata. In short: innate, unconscious but direct cognitive access to the linguistic is necessary for the acquisition of CSPRs, while CSPRs and E-perceptible M-representations are necessary for the acquisition of conscious but indirect (because conventionally M-representational) cognitive access to the linguistic.

Note that the more obvious uses of M-representations in communication and thought may be characterised in terms of their basic function as a conscious cognitive route to their linguistic representata. With regard to communication, the M-representational could be said to provide an individual with conscious cognitive access to linguistic expressions whose M-representations he E-perceives, and to linguistic expressions whose M-representations he E-produces for another individual to E-perceive. As for the use of M-representations in thought, the M-representational could be said to provide the individual with conscious cognitive access to linguistic expressions whose M-representations he uses introspectively, as it were -- either by E-producing them (i.e. "speaking to himself"), or by mentally entertaining and manipulating mental representations of them (i.e. thinking in the medium of internalised E-physical M-representations).

My view is that there are no significant differences between all the above uses of the M-representational. The remarks of Carston and Sperber and Wilson about the advantages of "linguistic codes" (in my terms, Conventional Systems of Physical Representations) for communication -- namely, that

> ... the most striking feature of linguistic communication is that it can achieve a degree of precision and complexity rarely achieved in non-verbal communication. (Sperber and Wilson 1986a/ 95: 174)
>
> A linguistic system... allows us to achieve a degree of explicitness, clarity and abstractness not possible in non-verbal communication... (Carston 1998: 56)

-- could as well be applied to the role of M-representations in thought. What distinguishes the sort of thought which involves the use of M-representations is that the thinker has much more conscious -- in fact self-conscious -- control over the course of it. This is because it has a medium which the thinker is able to consciously access and manipulate. Furthermore, the medium is in the form of a recursive system of what is therefore an infinite number of M-representations, the most simple and less complex of which are in specific M-representational relations to certain syntactico-semantic objects (i.e. linguistic expressions) which appear to constitute units of thought smaller and more basic than those which an individual would otherwise have access to. A briefer (though probably more cryptic) way of putting what I have just written is: the sort of thought which involves internalised E-physical M-representations is more "explicit", "precise" and "complex" because it is -- in virtue of having an M-representational medium -- more perspicuous than any other sort of thought as far as *conceptual compositionality* is concerned.

Another question about the use of the M-representational is whether it is its role in communication or in thought that is primary. Since the CSPR, whose acquisition goes hand in hand with the acquisition of conscious cognitive access to the linguistic, tends to be regarded as an essentially sociocultural phenomenon, to do with people interacting with one another, I am inclined to assume (with Dummett) that it is the use of M-representations in communication that is their primary role. On the other hand it is not inconceivable that even the smallest and most rudimentary system of Saussurean signs,

with any available E-physical phenomena serving as the signifiers or M-representans, might constitute a CSPR; and that such a CSPR may be acquired by someone isolated from society. Also, the very process of CSPR acquisition presupposes a pre-communicatory ability to internalise, and to attach conventional M-representational significance to, E-physical phenomena -- an ability which may be sufficient for an individual to think in an internalised E-physical medium. In short, I am suggesting that it may not be a matter of conceptual necessity that the conventional arises from the social, even if that is how it generally arises as a matter of empirical fact. This is an issue which requires further discussion and research.

5.2.2 M-representations, Carruthers and conscious thought

At this point I have to stress once again that, although E-physical M-representations of linguistic expressions provide conscious cognitive access to the linguistic, they *do not in themselves constitute conscious thoughts*. With regard to this, a comparison with Carruthers 1996 may prove useful. Carruthers claims that some -- possibly most --- kinds of conscious thought are constituted by what he calls "natural language sentences" (henceforth NLSs).

There are three points about his thesis which require stating and some elucidation. Firstly, his view of the linguistic is clearly externalist. His "natural languages" are "public languages", while his NLSs are "spoken/ heard/ written", and interchangeable with "utterances" and "inner speech/ dialogue". In short, the NLSs which he claims conscious thought involves are, in RC terms, actually E-physical M-representations of the linguistic.

Secondly, Carruthers' NLSs are not involved in *all* conscious thought. In his section 8.8 he argues that they are required for conscious thoughts about anything "beyond

immediately perceptible aspects of the spatial environment", and certainly for "complex" and "sophisticated" concepts like *personal identity* and *objective truth*.

Thirdly and most importantly, Carruthers' argument is that "natural languages" constitute not just the media or "representational systems" in which (some) conscious thoughts are formulated, but also the content of those thoughts. In his section 4.6, Carruthers asserts that to argue this successfully is to vindicate Chomsky's position on "natural languages" being as essential to cognition as it is to communication. Ironically, Carruthers' notion of NLSs as constitutive of (some) conscious thoughts seems much stronger than Chomsky's views on the subject. Carruthers believes that such NLSs should have a completely language-based semantics -- one that is independent of thought-based conceptual systems -- in order for those conscious thoughts which they constitute to be wholly, obviously "linguistic". But what Carruthers wants to avoid -- the notion of a linguistic semantics that is derived from the semantics of some pre- and non-"linguistic" cognitive system -- is precisely what Chomsky (not to mention relevance theorists) seems to endorse. As I pointed out in footnote 11 of my section 2.1.2, Chomsky maintains that there is an innate stock of concepts (1992a: 115, 116), and that I-language acquisition is merely the discovery of the "*labels*" of these innate concepts (ibid: 113, 116). (In RC terms, these "labels" are clearly E-physical M-representations.) Furthermore, Chomsky suggests that the innate concepts are "[drawn] from the resources of other genetically determined faculties of the mind" (ibid: 115). Even more explicitly and categorically, he states that

> *The a priori framework of human thought* within which language is acquired, provides necessary connections among concepts, reflected in connections of meaning among words, and more broadly, among expressions involving those words... (1992a: 114, my italics)

In any case, it seems to me that Carruthers does not manage to state his case coherently, let alone argue it successfully. For a start, he refers to "natural languages" as the primary *vehicle* of conscious thought (e.g. in his section 2.3), as a higher level of (obviously not M-) *representation* (e.g. section 8.1), and as being used in the *expression of thought* (e.g. section 3.3) -- as if NLSs and the conscious thoughts they are supposed to *be* were two separate things. Furthermore, just as Chomsky's "linguistic expressions" have to interface with the performance systems, and the information provided by Fodor's linguistic module has to be operated upon by the central processes, so Carruthers himself concedes (in section 8.2) that his NLSs must be accessed, selected and manipulated by a higher cognitive system like Gathercole and Baddeley's (1993) "central executive".[12]

Thus Carruthers' reasoning appears to proceed along these lines: being inherently and independently possessed of a semantics, NLSs can be accessed and manipulated by a higher "central executive" -- i.e. USED -- to express what conscious thoughts express, and can therefore constitute conscious thoughts. I have two objections to this. Firstly, if an NLS is a vehicle which has to be *used* to express/ become a conscious thought, then the NLS by itself is neither sufficient for, nor constitutive of, that conscious thought. Secondly, the natural language semantics which is supposed (by Carruthers) to constitute the propositional content of the NLS-involving conscious thoughts -- and in virtue of which the NLSs themselves are supposed to be at least as necessary an element of those conscious thoughts as the *use* of the NLS by the "central executive" -- is not intrinsic to the NLSs at all. The upshot of these objections is a sort of inversion of Carruthers' arguments: in my view, his NLSs are nothing more than E-physical phenomena; but being E-physical and therefore the most immediate and graspable/ manipulable objects of

---

[12] Actually, Carruthers presents Chomsky and Fodor as diametrically opposed in their respective views of *the role of the linguistic in cognition* -- and it is true that Chomsky believes that his "linguistic expressions" can be used in thought, while for Fodor thought is the exclusive business of the central systems and Mentalese (or "the language of thought"). My point here, on the other hand, is that their respective views on *how cognition bears upon the linguistic* do appear to coincide.

conscious thought, they can be internalised and used as a medium of conscious thought, and for that purpose are conventionally endowed with meanings (and referred to as "NLSs").

It follows from this that the propositional significance and conscious nature of thoughts which involve NLSs are *not constituted by the NLSs, but by the cognitive response to the NLSs*. Even Carruthers admits as much when he claims (in his section 7.1) that consciousness consists in an "accessibility-relation" to the NLS-vehicles of thoughts.[13] More generally, my point is that it is hardly controversial to claim that a great deal of conscious thought consists in the mental representing and manipulating of E-physical Saussurean signifiers. The E-physical and E-perceptible can be (consciously) used as signifiers or vehicles in conscious thinking (as well as in communication), precisely because they are *accessible to* conscious thought. It is for the same reason that they cannot themselves *constitute* conscious thoughts.

5.3 A naturalist view of the relation between the linguistic and the "language of thought": the linguistic *as* "language of thought"

When I say that Carruthers' NLSs are "signifiers or vehicles" I of course mean that they are non-natural M-representations of linguistic expressions, and that their "meanings" are constituted by their cognitive significance as M-representations of linguistic expressions. This cognitive significance arises, not only from the *M-representational nature of the M-representations*, but also from the *semantic nature of their linguistic M-representata*. So

---

[13] In his section 4.6 Carruthers does, however, equate an NLS's semantics with the contents, and -- astonishingly -- the NLS's *(internalised) E-physical form* with the *conscious state*, of the corresponding conscious thought. I.e he seems to be saying here that *the state of being conscious of E-physical U* is *constituted by U* (whether U is the non-natural E-physical vehicle of meaning M, or just any E-physical object). That seems to me downright wrong, not to mention bizarre.

far I have been discussing the E-physical vehicles of what is traditionally regarded as linguistic -- Carruthers' "natural languages", Sperber and Wilson's "linguistic codes", Chomsky's "I-languages". I have also been assuming, in the process, that the (literal, conventional) *meanings* of the M-representational vehicles ARE the M-represented *syntactico-semantic objects* (i.e. the linguistic expressions). In 5.3.1, for example, I took Mrs Malaprop's use of the E-physical {derangement} to literally mean *arrangement* to be the same as Mrs Malaprop's use of the E-physical {derangement} to M-represent 'arrangement'.

All this requires my system of notational conventions to undergo another adjustment. I have been using single quotes and italics to distinguish between linguistic expressions and their (structured) semantics. In the previous paragraph, however, I equated the contents of the single quotes (the linguistic expression 'arrangement') with what is italicised (its linguistic semantics *arrangement*). For the rest of this chapter, my view of linguistic expressions as vehicle-less syntactico-semantic objects will be reflected by my use of italics rather than single quotes to indicate linguistic expressions. In other words: to show that a structured concept like *arrangement* is in itself a linguistic expression, the italics will henceforth take over the role of the single quotes. So instead of distinguishing between the linguistic expression 'arrangement' and its semantics *arrangement*, I take *arrangement* to be the vehicle-less syntactico-semantic object which CONSTITUTES the linguistic expression, and which is M-represented by {arrangement}.

However, recall that the semantics of a linguistic expression as discussed in chapters 2 and 3 -- Chomsky's LF representation, the linguistically encoded logical form of Sperber and Wilson -- not only is separate from, but also falls short of, the truth-conditional, propositional meaning expressed through the use (even the literal use) of that linguistic expression. Hence the distinction between linguistic semantics and the truth-theoretic or "real" semantics of the central cognitive systems. Hence also the assumption that the former is inherited or drawn from (and somehow less than) the latter, and distinguished as

linguistic by its *labels* (Chomsky 1992a: 116) or by its being *encoded* by sentences or utterances (relevance theory). However, I concluded in chapter 3 that linguistic semantics, being logically related to the "real" semantics from which it is "drawn", has to be as truth-conditional as, and ontologically identical to, the "real" semantics of what is communicated or thought. And according to Burton-Roberts' representational conjecture (RC): labels, sentences and utterances are not linguistic but M-representational of the linguistic; Chomsky's I-languages and Sperber and Wilson's linguistic codes are Conventional Systems of Physical Representation (CSPRs); and linguistic expressions are exclusively syntactico-semantic objects. The question that remains is: what is the distinction -- if there is indeed a distinction -- between linguistic semantics and "real" semantics? In other words, what is the distinction between

> (i) the linguistic expression, with its structured linguistic semantics, which one intends a particular E-physical phenomenon to M-represent, and
>
> (ii) the "language of thought" (or "mentalese") expression, whose "real" semantics constitutes the proposition or propositions which one intends to express through the use of the M-representation of the linguistic expression?

In this section I argue that, with the RC's elimination of the conventional (Saussurean) vehicle-meaning relation from the linguistic, linguistic semantics as a whole is at least partly constitutive of "real" semantics. I also argue that the traditional notion of "language" *USE* -- in communication or thought -- is inapplicable to the truly I-linguistic (as conceived of in RC terms) and to the "language of thought" (henceforth LOT). Finally, I discuss the implications of my claims for theories of thought and consciousness.

According to Fodor (1975, 1987), there is a fundamental, even physiological distinction between the LOT and what are generally referred to as "natural languages" (e.g. the relevance-theoretic linguistic codes, Chomsky's I-languages). This distinction is supposed to be coextensive with that between the central systems and "linguistic" input systems.[14] As I mentioned in chapter 3, the notion of the LOT is presupposed by the computational-representational theory of thought which Fodor and the relevance theorists subscribe to, and according to which thought has to do with the computations of the central cognitive processes over the formulae of some *internal language*. This internal language or LOT is also presupposed by my (and of course Carston's) discussion of a "real" semantics -- i.e. the semantics of what humans express in thought or communication, as opposed to the semantics of so-called "natural languages". The very notion of a "real" *semantics* indicates that there is some language which it is the semantics of, just as a linguistic semantics is the semantics of a "natural language".

Fodor offers two relatively independent reasons for an LOT separate from "natural languages". Firstly, he points out that "there are non-verbal organisms that think" (1975: 56). "Infrahuman organisms" and "preverbal children", he argues (ibid), are capable of mental processes involved in considered action, concept learning and perceptual integration. If such mental processes are computational, then there must be a *representational system* for *representing* what they compute over. Note that "REPRESENTATION" or "representational system" in the general Fodorean (and relevance-theoretic) sense is quite different from, and much vaguer than, Burton-Roberts' "M-REPRESENTATION" or "Conventional System of Physical Representation". It is not clear if the Fodorean "to represent" means *to symbolise* or *to embody*. To be more precise, it is not clear if Fodorean "representation" corresponds to BR's "M-

---

[14] It does not seem quite right to distinguish between the LOT and "languages" (e.g. relevance-theoretic codes) by referring to the latter as "NATURAL languages". After all, the LOT is at least as natural as, if not more natural than, "natural languages". More on this below.

representation" or "C(onsitutive)-representation" (see section 5.1). On the one hand, Fodorean representations appear to be the *vehicles* of what the mental processes compute over. On the other hand, there is Fodor's use of "representation" in the following passage:

> It is nomologically necessary that the internal *representation* of the grammar (or, equivalently for those purposes, the internally *represented* grammar) is causally implicated in communication exchanges between speakers and hearers... (1985: 149, my italics)

That it is the "internal representation" of the grammar which is "causally implicated in communication exchanges" appears to indicate that the "internal representation" of the grammar is *constitutive of* (rather than the vehicle or medium for) the grammar.

The second independent reason given by Fodor for a distinction between "natural languages" and the LOT has to do with the issue of first language (i.e. "natural language") acquisition. Fodor argues that

> ... you cannot learn a language whose terms express semantic properties not expressed by the terms of some language [i.e. the LOT] you are already able to use. (ibid)

As I mentioned in the previous section, this appears to be consistent with Chomsky's suggestion that "language" is acquired within an "a priori framework of human thought" (1992a: 114).

I must stress that, although the LOT is conceived by Fodor and relevance theorists as distinct from "natural languages", it is as much of a "*(Fodorean) representational system*" as any "natural language". That the LOT and "natural languages" are in fact "languages" in exactly the same sense is a point that is explicitly made by Sperber and Wilson (1986a/95: 173-4). For Sperber and Wilson, languages are tools for both communication and thought -- the main difference between the LOT and "natural languages" is that the former

is an *internal language* used for "the processing and memorising of information" (ibid: 173) while the latter is an *external language* used for communication.

Furthermore, Fodor appears to believe -- with Chomsky (1992a: 113-6) and Carston (1998: 64-5) -- that a "natural language" shares (at least a substantial part of) its semantics with the LOT, as his 1975 (: 109-15) concept of "messages" clearly demonstrates. On the one hand, what Fodor calls "messages" constitute "linguistically carried information" (ibid: 111) -- i.e. what hearers recover or speakers convey during verbal communication by mapping wave forms to messages or vice versa. On the other hand, messages are also "formulae in the language of thought", since linguistically carried information is processed by the central systems, and the central systems are supposed to operate only over internal language formulae (ibid: 115).

However, Fodor claims that the nature of the LOT must be different from that of "natural languages", in order for the problem of infinite regress to be circumvented. As described by Fodor himself (ibid: 65), the problem is: if the understanding of a language predicate involves representing the extension of the predicate in some metalanguage, then presumably the understanding of a metalanguage predicate would involve representing its truth-conditions in some meta-metalanguage, and so on. For Dennett (1969: 86-7), the problem of infinite regress arises from the very notion of a "language of thought" -- and, more specifically, from the notion of "sentence-analogues in the brain" (ibid: 86). Dennett argues (ibid) that it is the "syntactic features" of a "sentence", and the "reading off" and analysis of these syntactic features, which determine the "function" which the "sentence" has "within a particular system". The problem is that a syntactically analysable "sentence" of the LOT would be syntactically analysed by a mechanism which may itself have syntactically analysable internal states. Thus Dennett concludes that "the regress must end

eventually with some systems which store, transmit, and process information in non-syntactic form".[15]

But according to Fodor (1975: 65-8), there is no such regress because the predicates of the LOT are *used* and *understood* differently from the expressions of "natural languages". Whereas the (Fodorean) representation of the extension of a "natural language" predicate is used/ understood in virtue of being given in some other language which we already understand (i.e. the LOT), the understanding of an LOT predicate "might just be that one's use of the predicate is always in fact conformable to the truth rule" (ibid: 65). Or, an individual is "so constructed that its use of the [LOT] predicate (e.g., in computations) comport with the conditions that such a representation would specify" (ibid: 66).

Fodor provides an analogy in the form of computers' machine (as opposed to input/ output) language. Formulae in the machine language

> ... correspond directly to computationally relevant physical states and operations of the machine: The physics of the machine thus guarantees that the sequences of states and operations it runs through in the course of its computations respect the semantics constraints on formulae in its internal language. (ibid)

---

[15] From Dennett 1969 it would appear that Dennett's notion of *syntax* is inextricably linked with the notion of *morphology*. For example, he writes of "brain-*word* tokens" (1969: 87, my italics) which are, presumably, the units from which "sentence-analogues in the brain" are constructed. He also refers (ibid: 86) to the syntactic structure of a "sentence token" as "the ordering of the word tokens". In the context of the LOT, it is not clear what this "*ordering*" entails, and how it is different from the sort of spatial or temporal order which applies to E-physical utterances (i.e. M-representations) of linguistic expressions. Or rather, it is not clear what Dennett thinks the morphosyntactic aspects of an obviously non-phonological LOT involve. For the RC, of course, even *linguistic* syntax is *semantic* structure; and the units from which linguistic expressions are constructed are not "words" in the morphophonological sense, but *concepts* (or what Chomsky might call "logical forms"). But whatever Dennett really means by "syntactic" or "non-syntactic", it is quite clear that he regards the "syntax" of a "sentence" as *functional* -- as a sort of structural complexity whose *raison d'être* is to be analysed as part of the process of "reading" and "understanding" the "sentence".

Therefore the "truth definition" of the machine language is not a function that maps machine language formulae onto formulae of an internal code, but a set of engineering principles which guarantees the "direct correspondence" between formulae and physical states/ operations. Similarly, the formulae of Fodor's LOT are "assigned" to neurological states of the organism

> ... in such a fashion that (some, at least) of the sequences of states that are causally implicated in the production of behaviour can be interpreted as computations which have appropriate descriptions of the behaviour as their 'last line'. (ibid: 73)

In other words, there is no further regress for Fodor's LOT because the LOT formulae are "interpretations" or "descriptions" which directly correspond to (physical) neurological events with causal properties. As Fodor stresses (ibid: 74 fn 14), a certain LOT formula is not *read as* commands for activating a particular system E -- it *IS* the physical event of type P which is causally sufficient for activating E, *interpreted as* an internal code message with E-activating properties.

However, in order for an LOT formula to actually constitute a particular mental state of an organism, it has to be in a particular computational relation to that organism (ibid: 75-7, 198). In Fodor's words: "To have a certain propositional attitude is to be in a certain relation to an internal [Fodorean] representation" (ibid: 198). For each propositional attitude there is a relation between an LOT formula and the organism which is "nomologically necessary for (or nomologically identical to)" having the propositional attitude. I assume that being in a computational relation to an internal (Fodorean) representation is to have one's central processes operating on the internal representation in a certain way.

Fodor also claims that humans can *use* LOT formulae (ibid: 172-4) -- i.e. that internal (Fodorean) representations can themselves be represented for certain cognitive purposes -

- and that "the intelligent management of internal representations" is "a fundamental and pervasive feature of higher cognitive processes" (ibid: 164) He argues that humans can control what Fodorean representations get assigned to what inputs and outputs, and that an individual would have to be able to represent the different ways of representing inputs and outputs in order to choose between them. Another way in which LOT formula may be manipulated, Fodor suggests, is when an individual incorporates ("external") natural language terms into the ("internal") LOT:

> It does not follow that for every natural language predicate *that can be entertained* there is an *entertainable* predicate of the internal code. It is no news that single items in the vocabulary of a natural language may encode concepts of extreme sophistication and complexity. If terms of the natural language can become incorporated into the computational system by something like a process of abbreviatory definition, then it is quite conceivable that learning a natural language may increase the complexity of the thoughts that we can think. (ibid: 85)

The question is, are these notions of *entertaining/ using LOT formulae* and *adding new representations to the LOT* compatible with the idea of representations which correspond directly, and are in a causal relation, to neurological states? According to Fodor's own description of the LOT, LOT formulae appear to be "neurally hardwired" -- each formula *is* a particular neural structure, and its semantic content is directly dependent on, and possibly constituted by, its neurological aspects. Thus LOT formulae are (at least partially) constitutive of the mental states which have to do with the use of Fodorean representations. My argument is that new LOT formulae cannot be created if the LOT is regarded as part of the neural architecture; and, more importantly, that LOT formulae cannot be *used* since they are themselves involved in the use of representations.

Basically, this is the problem of infinite regress. As I mentioned above, Fodor 1975 and Dennett 1969 approach it from the angle of how "language predicates" are understood: via previously understood metalanguage predicates, the understanding of

which is in turn dependent on a prior understanding of meta-metalanguage predicates, and so on. I argue that infinite regress also arises from the notion of *use* -- or, more specifically, the use of a "language" as a vehicle or Fodorean representational system in thought and communication. This presents a problem for Fodor as well as for Sperber and Wilson. The use of Sperber and Wilson's external languages (i.e. "natural languages") as a medium for communication is effected by the cognitive processes of the central systems, and also by the use of the internal language as a medium for this cognitive processing. But then the *use* of the internal language as a medium for cognitive processing would involve further cognitive processing, presumably in an even more internal language. Thus it would appear that the notion of the LOT or internal language is simply incompatible with the notion of use -- and, by extension, with the notion of *vehicles that have to be used.* In other words, the LOT cannot be a Fodorean representational system like a "natural language", consisting of Fodorean representations or vehicles which are used in virtue of the "meanings" to which they are conventionally related.

In my view, the infinite regress problem originates in the fundamentally *externalist* nature of the computational-representational theory of thought, and of the Fodorean (and relevance-theoretic) claim that the LOT is a usable (Fodorean) representational system like "external/ natural languages".[16] These "external/ natural languages" are systems of E-physical (but actually non-natural) vehicles which are selected and manipulated, decoded and interpreted[17] -- i.e. USED -- in external behaviour (and especially in communication). However, E-physical vehicles can also be mentally *M*-represented and mentally used, and it is this (conscious) "processing" of E-physical Fodorean representations which constitutes the only sort of mental activity to which humans have conscious access. As a

---

[16] Even more externalist in nature is Carruthers' claim that *A PARTICULAR* LOT -- as opposed to *THE* LOT of Fodor -- is in fact constituted by an externalist "natural language".

[17] "Interpreted" is used rather loosely here. The fact is that E-physical objects -- including those which are generally referred to as "*utterances*" -- are intrinsically meaningless, and therefore cannot actually be said to be "interpreted".

result, the computational-representational properties of this externally-orientated use of E-physical vehicles tend to be projected by humans onto those other cognitive processes which are inaccessible to consciousness.

One of the representational conjecture's most important points is that the relation between the intrinsically E-physical and the intrinsically I-cognitive is merely M-representational, emphasising the fact that there is absolutely no necessity -- indeed, that it is *downright impossible* -- for the E-physical and the processes in which it is involved (i.e. what we can consciously access) to be anything like the I-cognitive (i.e. what we cannot consciously access). Thus my argument is that a "natural language" is a usable but non-linguistic and externally-orientated Conventional System of Physical (M-) Representation (CSPR). In contrast, neither the wholly internal and cognitive linguistic representatum nor the wholly internal and cognitive LOT is designed for use. Within this internalist/ naturalist framework (as conceived by Chomsky and Burton-Roberts), the only thing that is usable is the E-physical M-representation with its conventional relation to some linguistic expression. The linguistic, being a natural state of mind/ brain, precludes the E-physical, the use of the E-physical to M-represent the linguistic, and the non-natural relation between E-physical vehicles and their meanings. The same can be said of the LOT, even independently of the fact that the idea of using the LOT leads to infinite regress.

With the whole idea of a vehicle-meaning relation eliminated from both the linguistic and the LOT, the consequences are (I suggest) as follows. Firstly, there would be nothing to distinguish vehicle-less linguistic expressions from the equally vehicle-less LOT formula. More specifically, the linguistic semantics -- which is anyway supposed (by Carston (1998: 64-5), for example) to be "inherited" from "real" semantics -- *would not have the phonological (or other E-instantiable) "vehicles" to distinguish it from the LOT's "real" semantics*. Hence:

(5) The vehicle-less linguistic would be at least partly constitutive of -- rather than the medium for (some of the "real" semantics of) -- the vehicle-less LOT.

Secondly: whereas "natural languages" and "linguistic" input systems are merely M-representational of the linguistic, both the linguistic and the LOT are directly, immediately *of the central cognitive systems*. And whereas "natural languages" are *used by the central cognitive systems* to M-represent the linguistic,

(6) Both the vehicle-less LOT and the vehicle-less linguistic are (at least partly) constitutive of -- NOT the medium for, or used by -- the central cognitive processes (including those which effect the M-representational use of "natural languages").

And finally:

(7) On the assumption that (i) the central cognitive processes give rise to consciousness and conscious thought, and that (ii) the linguistic and the LOT are at least partly constitutive of the central cognitive processes: *linguistic expressions and LOT formulae would themselves be inaccessible to conscious thought.*

Under the computational-representational view and the associated notions of *vehicles* and *use*, thought is unavoidably seen as consisting of two separate parts: (i) the Fodorean representation, which tends to be regarded as a "thought" in itself (in Carruthers 1996 and Sperber and Wilson 1986/95 (: 2), for example); and (ii) the cognitive processes which compute over the Fodorean representation. In contrast, (5) and (6) make the following points: the *Fodorean representations which are "computed over" by the cognitive processes* are *neither natural nor intrinsically I-cognitive*, while *the truly natural and I-*

cognitive linguistic expressions and LOT formulae are *not computed over by cognitive processes*. Instead, I suggest, the Fodorean representations are (internalised) non-natural M-representations of linguistic expressions, while linguistic expressions and LOT formulae are (at least partly) constitutive of the central cognitive processes. My argument, therefore, is that it is (ii) and (ii) alone -- including the linguistic and LOT constituents of (ii) -- which thought consists of.

However, I am not denying that there are more ways than one of thinking. Nor am I denying that at least one way leads to, and then proceeds via, *computations over vehicles or Fodorean representations (i.e. the use of M-representations of linguistic expressions)*. What I am suggesting is that it is the computations, not the vehicles or Fodorean representations, which constitute thought; and, furthermore, that vehicle-less linguistic expressions and LOT formulae are (at least partly) constitutive of such computations.

The linguistic and the LOT are also, I argue, (at least partly) constitutive of the sort of thought which has nothing whatsoever to do with M-representations and their use. This vehicle-less sort of thought appears to be as involuntary and as inaccessible to consciousness as the digesting of food by the alimentary canal, or the circulation of blood by the circulatory system. Perhaps it is less likely to be accessible to consciousness than the workings of the stomach or the heart, in virtue of *itself being constitutive of consciousness*.

In my view, to be conscious is simply to have a mental as well as a physical life -- to be able to be *aware of*, rather than just to be *impinged upon*. I am obviously in no position to discuss evolutionary neurology, and can only assume that consciousness evolved together with the central nervous system; that any organism with a central nervous system has the capacity for consciousness; and that consciousness varies in degree and nature according to the different central nervous systems of different species.[18]

---

[18] For example, what are conceptualised as *things* by humans may be regarded very differently by, say, dogs.

More controversially, I make the following claims about the relation between consciousness and the linguistic. What Chomsky or Carruthers calls "natural languages" (or what Burton-Roberts calls systems of physical M-representations) are not a prerequisite for consciousness. However, I argue that having *concepts* -- IN THE FORM OF VEHICLE-LESS, SYNTACTICO-SEMANTIC I-LINGUISTIC EXPRESSIONS AND LOT FORMULAE -- is both necessary and sufficient for an organism to be conscious. Thus not only am I arguing that

(8) To be minimally conscious is to (involuntarily, naturally) have concepts,

but also that

(9) Any organism with a central nervous system is in possession of some form of the linguistic (and thereby capable of conceptualising, and of being conscious).

What I have written above on the topic of consciousness is obvously far too sketchy and speculative to serve as anything other than an indication of some of the wider issues and questions raised by the representational conjecture (RC). These issues and questions require far more investigation than I am presently capable of. My purpose here is only to point out that there *are* such issues and questions, and that they have to do with what I perceive to be the conceptual (and probably empirical) overlapping of the linguistic, the LOT, thought and consciousness.

I have suggested that minimal consciousness in an organism is equivalent to the simplest, most fundamental form of thought, the sort of thinking which is as essential and involuntary as the workings of the alimentary canal or the circulatory system. This fundamental and involuntary sort of thought is (I argue) the starting point of all other, more complex and deliberate sorts of thought (including those involving the conscious use

of M-representations of linguistic expressions). It is also what all these other sorts of thought can be reduced to. Most importantly, it arises from cognitive processes which are (at least partly) constituted by vehicle-less, syntactico-semantic linguistic expressions and LOT formulae. It follows from this that the possession of the linguistic -- the natural, unique, austere and completely I-cognitive Language of Burton-Roberts' naturalist view -- is crucial to the conscious state, and to the fundamental sort of thought which the conscious state entails. The conscious state, in turn, is the crucial element of the cognitive processes which effect the use of E-physical phenomena to M-represent linguistic expressions, as well as the internalisation and internal use of the M-representations.

Note that the phrase "*conscious thought*" is rendered ambiguous (or rather, M-representationally vague) by my distinction between the fundamental, vehicle-less sort of thought, and the sort of thought which involves the conscious use of M-representations. If "thought" refers to the former, i.e. the involuntary mental activity which *constitutes consciousness*, then "conscious thought" would not only be a tautology, but would also refer to a process which is actually *inaccessible to consciousess*. On the other hand, "thought" may be described as "conscious" in virtue of being the sort of thought which is (partly and indirectly) accessible to consciousness -- i.e. the sort of thought which involves the conscious use of E-physical M-representations of linguistic expressions.

Once again, a comparison with Carruthers 1996 may serve to highlight some of my own points. My argument is that the linguistic is an integral part of conscious thought (in both senses), but not in the way Carruthers thinks it is. For Carruthers, "conscious thought" is constituted by the "natural language sentences" (NLSs) which are accessed, selected and manipulated by the "central executive". I, on the other hand, claim that the sort of NLS-involving conscious thought which Carruthers refers to is constituted, not by the NLSs, but by the "central executive"'s accessing, selecting and manipulating of the NLSs. According to the RC, Carruthers' NLSs are in fact neither natural nor linguistic.

Instead, they are mental (M-)representations of non-linguistic, E-physical M-representations of I-linguistic expressions.

Where the truly linguistic comes in, I argue, is at the level of the "central executive": together with the rest of the LOT, it at least partly constitutes the central cognitive processes which drive all forms of thought. In other words: whereas Fodorean representations are detachable from the computations which compute over them, the linguistic and the LOT are much more closely related (and probably inextricably linked) to the central processes. There are no intervening vehicles: linguistic expressions and LOT formulae consist only of what the vehicles are supposed to carry (or what Fodorean representations are supposed to be Fodorean representations of), and are therefore purely *conceptual* (or syntactico-semantic). As I have suggested, it is the capacity to have concepts which appears to distinguish the conscious state, and all thought could be said to be the result of cognitive processes that incorporate (rather than compute over) vehicle-less, purely conceptual linguistic expressions and LOT formulae.

I stress once again that thought need not -- though it may -- involve E-instantiable vehicles. At its most basic, it appears to consist in the consciousness of things in the external world *as things* (and as particular kinds of things), the capacity to have the concept *thing*, and the capacity simply to have concepts. If the linguistic is regarded as at least partly constituting these concepts, then there can be no distinction between BR's notions of "significance" and "linguistic sense" (i.e. linguistic semantics):

> A wide variety of phenomena have *significance* (red skies at night, smoke in the engine room, the lack of phone calls during the past week, the boss's silent exit from the room, (my placing) 'a' at the end of an essay) without being meaningful (sense bearing) in the favoured (linguistic) sense. The letter *t* (e.g. its occurrence in *the*) is significant, but in no sense meaningful or sense-bearing. (1994: 208 fn 7)

Smoke in the engine room is, of course, an E-phenomenon with a natural relation to the E-phenomenon of a fire in the engine room. The same goes for the lack of phone calls during the past week and its relation to, say, one's friend having departed for Australia. The *significance* of smoke in the engine room as an indication of a fire in the engine room, on the other hand, is an *I*-phenomenon. In other words, smoke in the engine room can only be said to have significance in the context of *being thought about*. My argument is that this "significance" of smoke in the engine room *IS* IN FACT "LINGUISTIC SENSE". Or rather, it is the CONCEPT of *smoke in the engine room arising from a fire in the engine room*, as constituted by the relevant vehicle-less linguistic expression. In my view, even being conscious of (as opposed to having unconscious nervous reactions to) smoke in the engine room -- *as nothing more than smoke in the engine room* -- involves the vehicle-less linguistic expression *smoke in the engine room*.

The sort of thinking which constitutes consciousness is in turn the basis of the conscious use of E-instantiable vehicles/ M-representations. Furthermore, the more complex sort of thought which has to do with the accessing, selecting and manipulating of M-representations of linguistic expressions could be regarded as an enrichment of the fundamental kind of thought. In the case of humans, the apparently highest level of consciousness which they have reached is a matter, not just of (*unconsciously*) having concepts, but also of having some sort of (*conscious*) idea of the concept of *concept*, and some (*conscious*) idea that one has concepts (including the concept *concept*).

As I claimed above, the process of consciousness-creating conceptualisation is itself an *unconscious* process. The *consciousness* of having the concept of *concept* (and the consciousness of having the concept that one has concepts) appears to be facilitated by the acquisition of conscious, reflexive cognitive access -- via the (conscious) use of E-physical M-representations -- to some of the conceptual constituents of the cognitive processes which underlie thought and consciousness. I agree with Carruther's point (1996,

section 8.2) about the "contents... of operations" of the "central executive"[19] -- which, within the naturalist and RC framework, include the linguistic and the LOT -- being in themselves *non-conscious*. But I would add that humans seem to have the unique ability to intentionally use E-physical objects to M-represent, and therefore consciously access, some of the "contents" of the "central executive". This access, however, is only M-representational, and therefore far from complete. So in a sense the "highest" level of consciousness is a somewhat compromised development -- reflexivity at the M-representational expense of transparency.

[19] What Carruthers actually says is that the contents and the *mode* of the operations of the "central executive" are unconscious. On the assumption that it is equivalent to my "vehicle", I reject the notion of central processes having a "mode".

6

The use of M-representations of linguistic expressions: some explanations, examples and questions

6 M-representation and the distinction between "semantics" and "pragmatics"

In this section I return to some of my arguments in 5.2.1 and 5.3:

> (i) Linguistic expressions generated by the unique, natural I-language are constituted solely by their structured semantics (LF).
>
> (ii) Linguistic expressions, as syntactico-semantic objects, are numbered (as it were) among the syntactico-semantic formulae of the LOT.
>
> (iii) Linguistic semantics is not -- as suggested in Chomsky 1992a (: 113-6), and Carston 1998 (: 64-5) -- inherited from, but is at least a part of, "real" semantics.

All this obviously has implications for the analysis of what is expressed during communication, and especially for the question of how exactly the linguistic is involved in communication. The general idea is that the speaker or writer produces an "utterance of a linguistic expression", and that this utterance tends to have "pragmatic content" and a "linguistic semantics". According to the representational conjecture (RC), however, the "utterance of a linguistic expression" is an intrinsically meaningless E-physical object, produced in aid of M-representing some linguistic expression. What the arguments of 5.2.1 add to this RC account of utterances is that any linguistic semantics involved is neither the "semantics *of*" the (intrinsically meaningless) utterance, nor even the "semantics *of*" the M-represented linguistic expression, but is in fact *constitutive of* the M-

represented linguistic expression. As a syntactico-semantic object, this linguistic expression is also as much a part of the LOT as any "pragmatic content" conveyed by the utterance. In other words, the M-represented linguistic expression and the LOT formula which constitutes the "pragmatic content" are at least *ontologically identical.*

My RC-based arguments also undermine the view that the distinction between "linguistic semantics" and "pragmatic content" is a *derivational* distinction. The general idea is that only the "linguistic semantics" is carried by the utterance, an E-physical "linguistic" vehicle, and that the E-perception of this "linguistic" vehicle will somehow lead the perceiver directly to the "linguistic semantics". For Sperber and Wilson (1986a/ 95), the process of recovering the "linguistic semantics" consists of the "automatic decoding" of the utterance (ibid: 177), while the "pragmatic content" is recovered via a "non-demonstrative inference process with free access to conceptual memory" (ibid: 65).

I, on the other hand, argue that the E-physical vehicle (i.e. the *M-representation of a linguistic expression*) and the vehicle-meaning (*M-representational*) relation are non-linguistic and non-natural. I also maintain that the "automatic" recovery of the "linguistic semantics" is only "automatic" insofar as it is *conditioned*, and, what is more, conditioned by *non-linguistic (usually sociocultural) circumstances and factors.* On the RC, an individual's "automatic decoding" of an utterance is in fact his recovery of the M-representatum of the utterance, governed in part by the particular Conventional System of Physical (M-)Representation (CSPR) which he has been *conditioned* to adhere to by his immediate social environment.

As Burton-Roberts and Carr (1999: 398-9) point out, the M-representation of linguistic expressions is such a fundamental and necessary human activity that the (M-representational) skills involved tend to be taken for granted. As a result of the necessity, and the "incessant practice" (BR and Carr 1999: 399) called for by the necessity, the M-representation of linguistic expressions (and the recovery of I-linguistic M-representata) might seem like an "automatic" process. Burton-Roberts and Carr observe:

> Were it as necessary for humans to play tennis as it is to externally M-rep[resent] the linguistic, we would find universal incessant early tennis practice and Wimbledon standard (or better) would be taken for granted. It might even seem mandatory and reflex-like. It is the explicit purpose of naval training in escape from submerged helicopters to make the required behaviour reflex-like. But this is true of anything we practise a lot (eg driving, touch-typing). It is the *necessity* that is interesting. (1999: 399 fn 16)

Note that the recovery of linguistic M-representata (or the production of E-physical objects for the purpose of M-representing linguistic expressions) only *seems* "mandatory and reflex-like", and seems so only if it is conditioned by "incessant practice". More importantly, the CSPR which an individual has been conditioned to use is not necessarily the CSPR he actually uses, or the CSPR used by the people he communicates with, on a particular occasion. It is in the nature of M-representation that what linguistic expressions are M-represented, and how they are M-represented -- in other words, what CSPR is used on a particular occasion by a particular individual -- is primarily a matter of *intention*. Consequently, there is nothing truly mandatory or "automatic" about the M-representation of linguistic expressions. The recovery of I-linguistic M-representata, in particular, is based on the assumption of (M-representational) intention, and is therefore chiefly a matter of fallible inference.

So my argument is this: it follows from the RC that there is simply no natural distinction, ontological or derivational, between "linguistic semantics" and "pragmatic content". With the E-physical vehicle (phonetic form) shifted to the domain of the (non-linguistic) M-representational, the remaining "linguistic semantics" has nothing to distinguish it from the LOT from which it is supposed to be derived. Furthermore, if the E-physical vehicle is in fact *non-naturally M-representational* of the linguistic expression constituted by the "linguistic semantics", then the so-called "linguistic decoding" of E-physical utterance phenomena (i.e. the recovery of linguistic M-representata) must be an

intentional, non-natural, M-representational process, and therefore neither linguistic nor automatic.

In this and the following sections I discuss various characteristics and strategies of the M-representation of linguistic expressions, and also use these M-representational characteristics and strategies to provide an alternative account of the non-natural distinction between "semantics" and "pragmatics", the equally non-natural distinction between literal and non-literal meaning, and communication in general. In Figure 1, I set out the different types of M-representation and the relations between them:

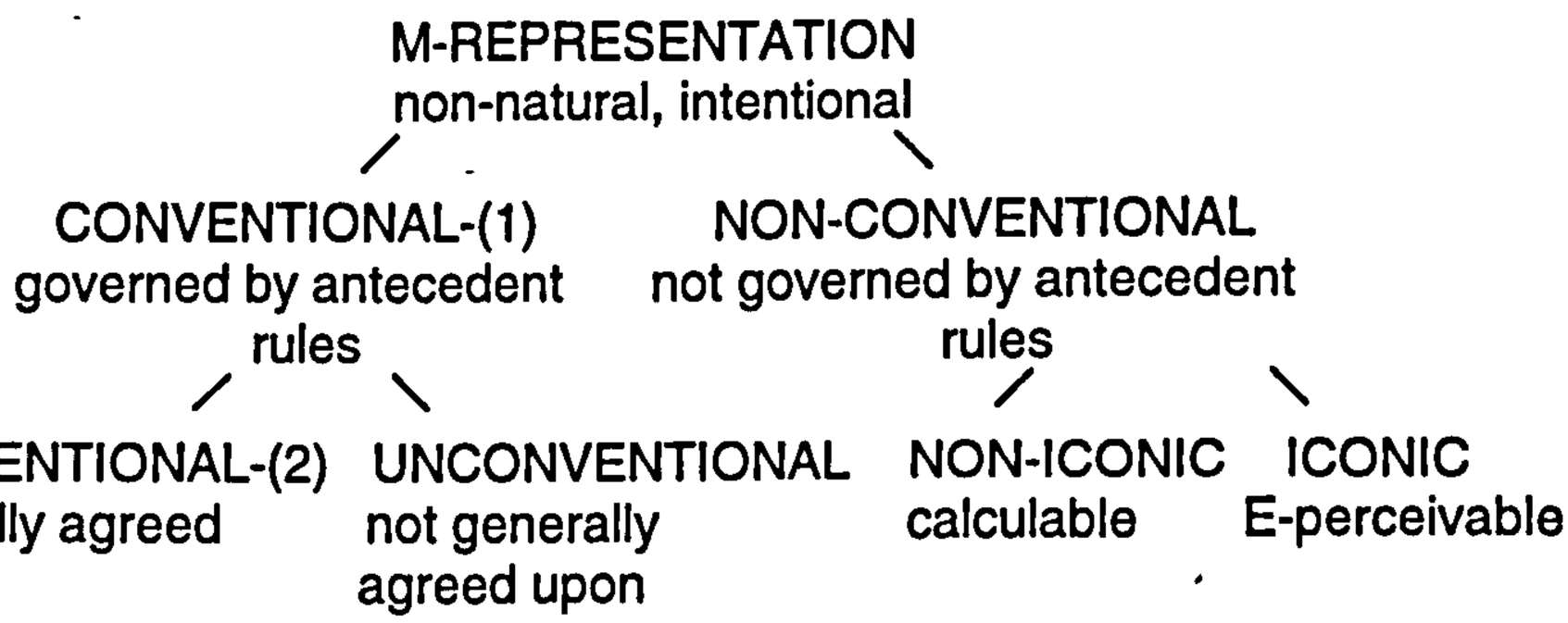

Fig. 1: Types of M-representation

M-representation, by definition non-natural and intentional, may or may not be constrained by antecedent rules. So far I have been using the term "CONVENTIONAL" to describe rule-governed M-representation. Thus an M-representational rule is necessarily conventional, simply in virtue of being a rule or convention. However, rules may also be conventional in the sense of being *generally agreed upon and publically sanctioned.* Henceforth I shall distinguish between "conventional" in the *rule-governed* sense and "conventional" in the *agreed upon* sense by labelling the first "conventional"

with (1), and the second "conventional" with (2). Note that only CONVENTIONAL-(1) M-representation can be CONVENTIONAL-(2) – i.e. that only rule-governed M-representation has the rules for people to agree upon. Thus "conventional-(2)" entails "conventional-(1)", but not vice versa. An example of M-representation which is conventional-(1) but not conventional-(2) is the M-representation of *arrangement* by Mrs Malaprop's {derangement}.

The opposite of conventional-(1) M-representation is NON-CONVENTIONAL M-representation. The non-conventional M-representation of *y* by *x* is intended (or assumed to be intended) to be noticed, arrived at, rather than specified by some rule. Non-conventional M-representation may be ICONIC or NON-ICONIC. As mentioned in 5.1, if *x* iconically M-represents *y*, it does so, not according to some rule or convention, but in virtue of being intended to share certain E-physical properties with *y*. The example I gave was of Magritte's painting of a pipe, which Magritte intended to share mainly visual properties with the pipe. Thus one can perceive the E-physical resemblance between *x* and *y*, and thereby discover what the iconic M-representatum *y* is. A non-iconic non-conventional M-representational relation between *x* and *y*, on the other hand, is one that is intended (or assumed to be intended), not to be perceived, but to be *calculated* or *inferred*.

The opposite of conventional-(2) M-representation is UNCONVENTIONAL M-representation. It is only the *agreed upon* aspect of conventional-(2) M-representation which falls within the scope of the prefix "un-". This means that, whereas an M-representational relation is *either* conventional-(1) *or* non-conventional, or perhaps partly conventional-(1) and partly non-conventional, the distinction between conventional-(2) and unconventional M-representation is a matter of degree. That is, one M-representational rule or system may be regarded as conventional-(2) -- or *more* conventional-(2), or *less* unconventional -- in virtue of having more adherents than another M-representational rule system. Alternatively, it may be regarded as

unconventional -- or *more* unconventional, or *less* conventional-(2) -- in virtue of having fewer adherents than another M-representational rule or system. Mrs Malaprop's {derangement} is conventional-(1) but not conventional-(2) – it is therefore an example of *unconventional* M-representation.

It follows from the sortal difference between (non-linguistic) M-representans and (linguistic) M-representatum that the E-physical M-representation of I-cognitive linguistic expressions may be conventional-(1) (i.e. conventional-(2) or unconventional), or non-iconically non-conventional -- but *never iconic*. That is, I assume the I-cognitive is never resembled by the E-physical (and the linguistic is never resembled by the non-linguistic). 6.1 focuses on the conventional-(2) and unconventional M-representation of linguistic expression, and 6.2, on the non-iconic non-conventional M-representation of linguistic expressions. In 6.3 I discuss "linguistic codes" and "literal meaning" in terms of the various types of M-representation described in 6.1 and 6.2. 6.4 and 6.4.1 are concerned with issues which I have so far not mentioned. 6.4 has to do with the (intentional) NON-M-REPRESENTATION of linguistic expressions, and its implications for the relation and/ or distinction between the linguistic and the LOT. 6.4.1 contains a brief M-representational account of REFERENCE, given in connection with a general discussion about how my radically internalist view of linguistic expressions and LOT formulae bears upon the perceptual/ cognitive interaction of humans with the external world.

6.1 Conventional-(2) and unconventional M-representation of linguistic expressions

The most conspicuous and recognisable sort of conventional-(2) M-representation is that which is constrained by the kind of CSPRs (conventional systems of physical representation) which constitute Burton-Roberts' "*particular languages*" (e.g. English,

Swahili). These CSPRs are "extralinguistic, culturally differentiated, conventional [i.e. conventional-(2)]" (BR and Carr 1999: 404). Dummett's 1986 "languages", it seems to me, are obviously conventional-(2) CSPRs. As CSPRs which are generally agreed upon and publically sanctioned, Dummett's "languages" would indeed be what he claims they are: sets of rules or conventions of *social practices*. And if the ideal Dummettian "language" is simply a particular set of M-representational conventions, endorsed by authority figures and publicly established, then it would follow that each member of a particular "language" community not only has an "imperfect" (Dummett 1986: 475), "partial, and partly erroneous" (ibid: 468) grasp of it, but deliberately deviates from it from time to time.[1]

That "particular languages" are extralinguistic, culturally differentiated, conventional-(2) CSPRs is further borne out by the CSPR-SPECIFIC phenomena which I shall call SECOND-ORDER and THIRD-ORDER (henceforth 2-o and 3-o) M-REPRESENTATION. Conventional-(2) first-order M-representation is simply a particular E-physical object conventionally-(2) M-representing a particular linguistic expression, and (as I argued in the previous section) thereby providing conscious (but incomplete) cognitive access to the LOT formula which that linguistic expression *is*. In contrast, conventional-(2) 2-o or 3-o M-representation involves the conscious accessing of an LOT formula (call it $LOTF_2$ or $LOTF_3$), not through an M-representation alone, but by way of an M-representation's conventional-(2) (first-order) M-representing of another LOT formula ($LOTF_1$). In that way, the M-representation can be regarded as conventionally-(2) M-representational of $LOTF_2$ or $LOTF_3$ in addition to -- and *in virtue of* -- conventionally-(2) M-representing $LOTF_1$. And the CSPR to which the M-representation belongs would specify that the M-representation is a conventional-(2) M-representation of $LOTF_1$ *as well as* of $LOTF_2$ or $LOTF_3$. In other words, the CSPR

---

[1] Thus the RC manages to reconcile a Dummetian account of "language*s*" with a Chomskian account of "Language" by stating the nature of the relation between them.

would include a 2-o or 3-o M-representational convention which mentions the corresponding first-order M-representational convention. [2]

To illustrate what I mean by conventional-(2) 2-o M-representation, consider English {take}. In the case of conventional-(2) 2-o M-representation, $LOTF_1$ and $LOTF_2$ are logically, conceptually related. What {take} primarily M-represents, according to the CSPR commonly known as English, appears to be the linguistic expression (i.e. $LOTF_1$) constituted by something like *take hold/ possession of*:[3]

(10)(a) She took the rolling pin and hit him with it.

(b) Someone has taken my pen.

But in certain other (M-representational) contexts, {take} in English appears to M-represent other concepts, like *do/ experience* --

(11) She wants to take a bath/ nap/ holiday.

-- and *bring* --

---

[2] Higher-order M-representation is by definition conventional-(1) (i.e. rule-governed). So 2-o and 3-o M-representational rules may feature in *unconventional* as well as conventional-(2) CSPRs.

[3] I am assuming that the first-order M-representatum of English {take} is *take hold/ possession of*, partly on the basis of my own beliefs about (or, in Dummett's words, my "partial, and partly erroneous" grasp of) the CSPR English, and partly because the entry for {take} in the Concise Oxford Dictionary is headed by the definitions "lay hold of" and "acquire, get possession of". However, it seems to me that each "English-speaking" individual has his own personal CSPR or CSPRs. These CSPRs are imperfectly derived from the public CSPR constructed by authority figures (e.g. the compilers of the COD), with the effect that they have rules in common with, but strictly speaking are absolutely distinct from, the publically established CSPR. Thus a particular individual's first-order and higher-order M-representational rules for {take} are not necessarily the same as those of, say, English as specified by the COD.

(12) She took the children home.[4]

Furthermore -- whereas the CSPR English appears to specify the first-order M-representation of *take hold/ possession of*, and the 2-o M-representation of *do/ experience* and *bring*, by the same M-representation {take} -- the CSPR known as Italian has separate M-representations for those concepts:

(10') prendere una penna

to take a pen

(11') fare un bagno

to do a bath

(12') portare a casa

to bring to home

Conversely, the Italian {fare} M-represents *do* -- as in (11') and

(13) fare i compiti

to do the homework

-- as well as *make*:

(14) fare rumore

to make noise

[4] In fact these are not the only 2-o conventional-(2) M-representata of {take}. What linguistic expressions, for example, does {take} conventionally-(2) M-represent in {take a bus}, {take some pills}, {take it that you agree with me}, etc?

As indicated, English deploys different M-representations -- {do} and {make} -- for the concepts M-represented by Italian {fare} in (13)-(14). A CSPR which does have the same M-representation for what in English is M-represented by {do} and {make} is Putonghua:

(15) zuo gongke

to do homework

(16) zuo zhifu

to make a uniform, or uniforms

However, according to the CSPR Putonghua, {zuo} only M-represents *do* and *make* in certain M-representational contexts. Unlike the Italian {fare}, it cannot be used with the M-representation of *noise* or *a bath*.

A more obvious form of conventional-(2) CSPR-specific 2-o M-representation has to do with M-representational constructions like (17)(a)-(d):

(17)(a) {to laugh in someone's face}

(b) {the man in the street}

(c) {a poor relation}

(d) {special pleading}

A person may know what the conventional-(2) first-order M-representata of (17)(a)-(d) are -- respectively (and approximately) (18)(a)-(b):

(18)(a) *to direct one's laughter straight towards someone's face*

(b) *some man who is situated in some street*

(c) *a relation which is poor*

(d) *pleading which is in some way special*

But if he does not also know the 2-o M-representational conventions of the CSPR English, he will not be aware that (17)(a)-(d) also have (conventional-(2), CSPR-determined) 2-o M-representata in the form of (respectively) (19)(a)-(d):

(19)(a) *to openly show scorn for someone*

(b) *the ordinary or average man*

(c) *an inferior member of a family or some group (of people or things)*

(d) *a fallacious and/ or unfair argument -- e.g. one which takes into account irrelevant details -- in support of someone or someone's point of view*

Coming now to 3-o M-representation: like 2-o M-representation, 3-o M-representation is CSPR-specific. For example, it is only in respect of the CSPR English that the M-representation {kick the bucket} is CONVENTIONALLY-(2) used to express *die* as well as *strike the bucket with one's foot*. The linguistic expression *die* is what I call the conventional-(2) 3-o M-representatum of {kick the bucket}. What distinguishes 3-o M-representation from 2-o M-representation is this. First-order and 2-o M-representata are, as I have pointed out, logically and conceptually related. To choose a man in the street is to choose arbitrarily, and hence to pick out someone who is representative of the community as a whole. Similarly, to guffaw at someone with whom one is face to face is a way of openly expressing scorn for him. In the case of 3-o M-representation, the first-order and 3-o M-representata are, semantically, totally distinct from one another -- as the {kick the bucket} example obviously demonstrates.

Falling somewhere between conventional-(2) 2-o and 3-o M-representation are examples such as {busman's holiday}. On the one hand, {busman's holiday} is not an example of 3-o M-representation, because the concept *leisure time spent doing what one usually does to earn a living* is semantically related to the idea of a busman spending his

holiday on a coach. On the other hand, this semantic relation may be missed by those for whom a busman going on holiday is not necessarily -- is in fact far from necessarily -- going on a coach trip.

Cases such as {busman's holiday} suggest that it is not always easy to decide whether a higher-order M-representation fulfils the criteria for 2-o or 3-o status. This has a great deal to do with how the M-representation is viewed by its users in connection with its first-order M-representatum. I stated that the distinction between 2-o and 3-o M-representation is that the first-order and 2-o M-representata of an M-representation are semantically or conceptually related, while the first-order and 3-o M-representata are not. However, it is clearly not the case that the 3-o M-representatum of, say, {pay through the nose} just happens, arbitrarily, to be *pay exorbitantly*. Thus a more accurate description of the distinction between 2-o and 3-o M-representation would be that the semantic relation between the first-order and 2-o M-representata of an M-representation is quite obvious, while that between the first-order and 3-o M-representata is no longer apparent. The accessing of 2-o or 3-o M-representatum via a particular first-order M-representation is a previously *non-conventional* custom which has been conventionalised, usually through familiarity and/ or frequent and widespread use. But whereas the semantic relation between first-order and 2-o M-representata is not obscured by the conventional relation, the rationale for the 3-o custom becomes opaque as the custom is conventionalised.

More importantly, whether or not the semantic relation is apparent (or at least immediately deducible) -- whether or not the rationale is transparent -- is an entirely non-linguistic, historical, sociocultural matter. In other words, the relative opacity or transparency of the relation between first-order and higher-order M-representata is determined by historical and sociocultural factors. For example, the connection between the conventional-(2) 3-o M-representatum (*pay exorbitantly*) and first-order M-representatum (which includes the concept *nose*) of the M-representation {pay through

the nose} has to do with Danes in the 9th century slitting the noses of people who refused to pay their taxes. This connection is no longer obvious because it is tied to a sociocultural/ historical circumstance -- i.e. having one's nose slit for not paying one's taxes -- which is not only very specific, but also no longer in existence. Much more general and persistent is the idea of a street being where someone representative of a community may be arbitrarily picked out. Consequently, the semantic relation between the conventional-(2) first-order and 2-o M-representata of {the man in the street} is still discernible. There may come a time, of course, when people do not walk in streets, or when streets have become obsolete. Such a situation could result in a shift from 2-o to 3-o status for the higher-order M-representatum of {the man in the street}.

Returning to {busman's holiday}, the reason why it is difficult to categorise may be that the rationale for using its first-order M-representatum to access the concept *leisure time spent doing what one usually does to earn a living* involves an association between holidays and coach trips which at present is not completely obscure, but in the past twenty years or so has become more and more tenuous. Let me stress once again that this is NOT a matter of "linguistic/ semantic change", but of historical changes in sociocultural environments (and in how sociocultural environments are perceived).

Note that conventional-(2) M-representata -- including conventional-(2) *higher-order* M-representata -- tend to be regarded as the "LITERAL MEANINGS" of their respective M-representations. *Take hold/ possession of, do/ experience* and *bring* are all accepted as the "literal meanings" of English {take}. The individual, simple (M-representational) constituents of a complex M-representation like {the man in the street} (i.e. {the}, {man}, {in}, etc) are generally seen as having "literal meanings"/ first-order M-representata that do not add up to the conventional-(2) higher-order M-representatum *the ordinary or average man*. But as far as the complex M-representation is concerned, the high-order M-representatum has as strong a claim on "literal" status as the first-order M-representatum. The conventional-(2) M-representans of a complex 3-o M-representation

like {pay through the nose}, in particular, is almost always used (or assumed to have been used) in its higher-order capacity, and almost never used to M-represent the conventional-(2) first-order M-representatum.

Furthermore, what constitutes the "literal meaning" of a conventional-(2) M-representation is by definition specified by the corresponding (first-order or higher-order) M-representational convention. A particular set of such conventional-(2) M-representational conventions -- i.e. a particular conventional-(2) CSPR -- constitutes a "particular language" like English or Italian. To "speak English", for example, is to grasp and use a particular set of conventional-(2) M-representational rules, including those which specify the respective first-order and higher-order M-representata of {take}, {the man in the street} and {pay through the nose}. Thus "particular languages" -- and the "literal meanings" of the "expressions" of "particular languages" -- are intentional, non-natural, extra-linguistic and socioculturally determined *M-representational* phenomena.

It has to be pointed out that, before they were conventionalised, the "literal meanings" constituted by conventional-(2) higher-order M-representata were regarded as "*metaphorical* (or at least non-literal) meanings". As mentioned above, the use of a first-order M-representation to access its higher-order M-representatum was formerly *not* specified by the corresponding conventional-(2) CSPR. Instead, it was a *non-conventional* process. That is to say, both higher-order M-representatum and its semantic relation with the first-order M-representatum were not learnt/ known (or assumed to be learnt/ known) together with the rest of the conventional-(2) CSPR, but intended to be *inferred*.

For example, the use of {pay through the nose} and its first-order M-representatum to access the concept *pay exorbitantly* appears to have been regarded as metaphorical in the mid-19th century. However, under some historical, sociocultural circumstance, the relation between {pay through the nose} and *pay exorbitantly* was conventionalised, and

incorporated into the relevant conventional-(2) CSPR. The result is that *pay exorbitantly* is now generally regarded as the "literal meaning" of {pay through the nose}.

In my view, the preceding discussion of conventional-(2) higher-order M-representation constitutes an M-representational account of what Davidson calls "dead metaphors" (1979: 35). According to Davidson, an example of a "dead metaphor" is {mouth} in the context of {the mouth of a river} or {the mouth of a bottle}:

> Once upon a time, I suppose, rivers and bottles did not, as they do now, literally have mouths... when "mouth" applied only metaphorically to bottles, the application made the hearer *notice* a likeness between animal and bottle openings... Once one has the present use of the word, with literal application to bottles, there is nothing left to notice. (ibid)

On the representational conjecture (RC), {mouth} in these M-representational contexts is a conventionalised 2-o M-representation (like {take} in {take it that you agree with me}), whose formerly non-conventional "metaphorical meanings" *place where a river enters the sea* and *the opening of a bottle* have become its conventional-(2) higher-order "literal meanings".[5]

I also claim that the problem of *idioms* -- a particular class of "dead metaphors" -- becomes trivially resolveable under my M-representational account of "dead metaphors". The problem is that idioms such as {kick the bucket} and {pay through the nose} are peculiar, not only to certain sociocultural contexts or groups, but also to certain "particular languages".[6] So the use of {kick the bucket} to express *die*, for example, appears to be governed by conventions that are "language"-specific, i.e. specific to the

---

[5] See Quine 1979 (:159) on the creating and "killing" of metaphors "at the growing edges of science and philosophy". See also Boyd 1979/ 93.

[6] Chomsky (1992a, 1995a) claims that English, Swahili, etc are not "I-languages", but non-linguistic "community norms". However, the same problem (i.e. the problem of the grammar of a "language" being unable to account for a construction specific to that "language") arises at the level of his "I-languages". For example, the use of {kick the bucket} to express *die* is specific to certain "I-languages" (Jones', Chomsky's) but not to others (e.g. Boris').

"particular language" English. Yet it is precisely the grammar of that "particular language" English which states that the "linguistic expressions" {kick} and {the bucket} respectively mean *strike with one's foot* and *the bucket*, and which cannot account for the particular "linguistic" construction {kick the bucket} also conventionally-(2) meaning *die*.

On the RC, however, {kick the bucket} is a conventional-(2) 3-o M-representation. In other words, it is simply *not* a linguistic expression with a semantics, but an E-physical object whose significance is merely M-representational. This significance, being M-representational, does not reside in the E-physical object, but in the mind of the individual who uses the E-physical object to M-represent the linguistic expression *die* on one occasion, and *strike the bucket with one's foot* on another. It does NOT follow from this, I stress, that either {kick the bucket}, or whatever {kick the bucket} M-represents, is *ambiguous*. The *M-representation* {kick the bucket} cannot be ambiguous, for it has no semantics. The M-representata *die* and *strike the bucket with one's foot* respectively and unequivocally constitute the semantics of -- and in fact individuate -- two distinct linguistic expressions.

Before I move on to unconventional M-representation, note that so far I have only discussed examples of complex M-representations whose conventional-(2) "literal meanings" are more often than not assumed to be their higher-order M-representata rather than their first-order M-representata. In the case of {beg the question}, however, it is debatable whether its "literal meaning" is its conventional-(2) higher-order M-representatum *presuppose the truth of an argument without arguing it*, or its apparently equally conventional-(2) first-order M-representatum *raise/ invite the question*. On the one hand, there are those for whom the "literal meaning" of {beg the question} is its higher-order M-representatum, in the same way that the higher-order M-representatum *pay exorbitantly* is regarded as the "literal meaning" of {pay through the nose}. On the other hand, there are those who seem to have no knowledge of the higher-order M-representational convention, and use {beg the question} "literally" to express its first-

order M-representatum *raise/ invite the question*. The important point is that the distinction between the first and second groups of people is not a linguistic distinction, but a matter of sociocultural M-representational variation. Thus the increasing number of people who use {beg the question} "literally" to express its first-order M-representatum seems to have something to do with a general tendency to acquire and use the less arbitrary and opaque M-representational conventions. Furthermore, the M-representatum specified by the less arbitrary first-order convention also happens to be a concept (*raise/ invite the question*) which is probably more frequently accessed/ expressed than the higher-order M-representatum (*presuppose the truth of an argument without arguing it*). In contrast, the first-order M-representatum of, say, {pay through the nose} involves a relation between the concepts *pay* and *nose* which may have been relevant to the 9th century nose-slitting Danes and their victims, but seems incongruous in this day and age. Not surprisingly, it is the higher-order rather than the first-order M-representatum of {pay through the nose} -- i.e. *pay exorbitantly* -- which at present is usually taken to be its "literal meaning".

The first-order M-representation of *die* by {die}, the 2-o M-representation of *the ordinary or average man* by {the man in the street} and the 3-o M-representation of *die* by {kick the bucket} are all specified by *English*, a CSPR which is generally regarded as *conventional-(2)*. But in fact the conventionality-(2) or unconventionality of any M-representational system used by two or more people is relative, and gradable. An M-representational system (intended or believed to be) used by just one individual is by definition *un*conventional. As far as I know, no one but myself uses {conventional-(1)} and {conventional-(2)} to M-represent, respectively, *rule-governed* and *agreed upon*. However, these M-representations are not as unconventional as those belonging to a private code/ CSPR used by someone who has no intention of sharing it with anyone else,

since I intend my M-representations {conventional-(1)} and {conventional-(2)} to be used by my reader, at least while he/ she is reading this chapter. Less straightforward is the use of the M-representation {M-representation}. For linguists in general, not to mention the community of people who use the CSPR known as English, {M-representation} is clearly an unconventional M-representation. For those who subscribe to Burton-Roberts' representational conjecture (RC), however, it is the most *conventional-(2)* way of M-representing the concept of M-representation.

Similarly, the particular uses of what Chomsky calls "terms of naturalistic inquiry" (1992a, 1995a) are conventional-(2) within the scientific community, but unconventional within wider contexts. In the first place, it follows from the RC that an "expression" like "acid" is -- in accordance with my own notational conventions -- more accurately presented as {acid}. Thus I reject Chomsky's assertion (ibid) that the individual's "I-linguistic sense" of {acid} is natural, while the chemist's "expert sense" of {acid} is not. As M-REPRESENTATIONS, both instances of {acid} are equally non-natural. But the so-called "I-linguistic sense" of {acid} could be said to be *more conventional-(2)* than the "expert sense", simply in virtue of the fact that the use of {acid} to M-represent *sour and/ or corrosive substance* is far more common and widespread than the "expert" use of {acid} to M-represent, say, *proton-donating atom or compound*.

The case of Mrs Malaprop presents another kind of complication. When another individual tries to recover the linguistic M-representata of Mrs Malaprop's utterances, that individual would have to adopt at least some of the M-representational conventions which Mrs Malaprop abides by. This is generally so in all cases of utterance interpretation. But before or after anyone has any verbal contact with her, Mrs Malaprop is clearly the only person who uses the M-representations which she uses. Such M-representations -- e.g. {epitaph} for *epigram* -- are therefore thoroughly unconventional. But whereas chemists (or at least most chemists) are aware of the relative unconventionality of their use of {acid}, Mrs Malaprop believes that her use of {epitaph} is usual, conventional-(2) and

generally agreed upon within the social group to which she belongs. Furthermore, Mrs Malaprop's unconventional M-representations, and Mrs Malaprop's belief that her unconventional M-representations are conventional-(2), are both rooted in the conventional-(2). According to the conventional-(2) CSPR of Mrs Malaprop's social group, {epitaph} is an M-representans (though not of *epigram*). {Epitaph} also shares certain E-physical properties with the conventional-(2) M-representans of *epigram*, {epigram}. According to the graphological CSPR, they begin with the same three symbols {e-p-i}; phonetically, they have two syllables in common. This is one of the reasons why Mrs Malaprop's M-representations, though unconventional, are not wholly unintelligible to those individuals who know only the conventional-(2) CSPR(s).

6.2 Non-iconic non-conventional M-representation of linguistic expressions

I have stated that conventional-(2) and unconventional CSPRs (conventional systems of physical representation) are by definition antecedent to any act of M-representation. However, as in the case of Mrs Malaprop's hearer/ reader, there is always the possibility of someone recovering what is (for him) *unconventionally* M-represented -- not just from the E-physical resemblance between the unconventional and the conventional, but also from other, more general contextual factors. For example, it is obvious from the two E-physical objects following Mrs Malaprop's {pineapple} -- {of} and {perfection} -- that Mrs Malaprop has used {pineapple} to (unconventionally) M-represent *pinnacle*. What distinguishes this from NON-ICONIC NON-CONVENTIONAL M-representation is that Mrs Malaprop did not intend the M-representatum *pinnacle* to be calculable. Instead, she assumed that it would be recovered on the sole basis of knowledge of a particular -- and, for her, conventional-(2) -- CSPR. The recovery of non-iconic non-conventional M-representata, on the other hand, is intended to be a reasoned process constrained by

contextual information and probably some general principle of communication and cognition like Sperber and Wilson's relevance theory. So what is non-iconically and non-conventionally M-represented is by definition never agreed upon in advance.

A fairly obvious case of non-iconic non-conventional M-representation has to do with the sort of utterances whose interpretation is supposed by Grice to involve the "Be orderly" submaxim of his Manner maxim -- e.g.

(20) Max got up and had breakfast.

For Carston (1998: 309), the hearer of (20) enriches the linguistically encoded logical form by adding a conceptual constituent: that of the temporal order of the two events (i.e. Max getting up and Max having breakfast). Carston claims that the additional conceptual constituent becomes part of the explicit content of the utterance. However, according to Sperber and Wilson (1986a/ 95), the proposition containing the concept of the temporal relation would be *less explicitly communicated* by (20) then it would be by

(21) Max got up and then had breakfast.

Sperber and Wilson define an "explicature" as

> ... a combination of linguistically encoded and contextually inferred conceptual features. (1986a/ 95: 182)

and state that

> The smaller the relative contribution of the contextual features, the more explicit the explicature will be, and inversely. (ibid)

(20) and (21) explicitly communicate the same proposition -- the proposition containing the concept of the temporal relation. But whereas (20) *implies* the concept of the temporal relation, (21) *encodes* it. Thus Sperber and Wilson would consider the explicature of (21) to be more explicit than that of (20).

On the basis of the representational conjecture (RC), a different account may be given of (20) and (21). I suggest that (20) and (21) express *Max got up and then had breakfast* with equal explicitness, because it could be argued that both (20) and (21) M-REPRESENT the linguistic expression constituted by the concept *Max got up and then had breakfast.* It is by *conventional-(2)* means, in accord with the CSPR English, that the linguistic expression is M-represented by (21) — and, in particular, that the concept of the temporal order of the events is M-represented by {then} in (21). In (20), however, the concept of temporal order seems to be M-represented -- *non-iconically and non-conventionally* -- by the *E-physical order* of the conventional-(2) M-representations {Max got up} and {[Max] had breakfast}.

Here is how the E-physical order of the conventional-(2) M-representations in (20) might (non-iconically and non-conventionally) M-represent the concept of the temporal relation. {Max got up} conventionally-(2) M-represents the linguistic expression which conceptually corresponds to the event of Max getting up, while {Max had breakfast} conventionally-(2) M-represents the linguistic expression which conceptually corresponds to the event of Max having breakfast. In (20), these two conventional-(2) M-representations are (intentionally) given a spatial order which resembles — and can therefore *iconically M-represent* — the temporal order of the events. Thus it would appear that the spatial order of the conventional-(2) M-representations *non-iconically and*

non-conventionally M-represents the CONCEPT of the events' temporal order, by way of *iconically M-representing* the temporal order of the events themselves.[7]

Let me stress that only the E-physical in this case -- the M-representans, the events (Max getting up and Max having breakfast) -- can be ordered, and thus iconically M-represent one another. The linguistic M-representatum, being I-cognitive, can have nothing to do with this E-physical order. What has to be borne in mind is the sortal difference between the I-cognitive and the E-physical, and the impossibility of there being any sort of resemblance between the two. It follows from this that the spatial order of the *E-physical* conventional-(2) M-representations may iconically M-represent the temporal order of the *E-physical* events, but does not -- and cannot -- iconically M-represent the *I-cognitive* CONCEPT of the temporal order. Rather, it M-represents the concept by non-iconic, non-conventional means -- namely, by iconically M-representing what the concept is the concept of.

The same principle -- of non-iconically M-representing a linguistic expression via the iconic M-representation of what the linguistic expression is the concept of -- appears to be central to the phenomenon and process of onomatopoeia. Take {mew}, as in

(22) The cat mewed.

In my view, {mew} *non-iconically* M-represents the *I-cognitive concept* of the sound which cats make, by physically resembling and thereby *iconically* M-representing *the actual E-physical sound which cats make*. There is, however, a difference between {mew} in (22) and the spatial order of the M-representations in (20). The M-representata of both are calculable -- can be worked out from the resemblance between E-physical M-

---

[7] Note that the spatial order in (20) is of graphological M-representations. In the case of (20)'s phonetic counterpart, there would be an even closer iconic resemblance between the temporal order of the events and the *temporal* order of the phonetic M-representations produced by the speaker.

representans and the E-physical phenomenon which the M-representatum is the concept of. But whereas the utterer of (20) expects the hearer/ reader to work it out, to perform the calculations, the person who uses {mew} in the context of (22) most probably believes that the M-representation of *sound made by cats* by {mew} is specified by the conventional-(2) CSPR English. In other words, what {mew} M-represents is calculable but no longer (intended to be) calculated, because there has been a superimposition of convention over calculation. So {mew} may once have been a non-iconic non-conventional M-representation of *sound made by cats*, but is now part of the conventional-(2) system of M-representation known as English.

I have been arguing that "external/ natural languages" like relevance-theoretic "linguistic codes" are in fact non-natural, conventional-(2) systems of E-physical M-representation -- that what is "*linguistically encoded*" is, on the RC, conventionally-(2) M-represented. From examples such as (20) it would appear that what relevance theorists regard as *pragmatially implied* -- in the case of (20), the concept of the temporal order of the events -- may also be M-represented (albeit non-conventionally). Indeed, since M-representations are by their (E-physical) nature linearly ordered, I am suggesting that the basis upon which the speaker selects the order of the M-representations which she utters -- e.g. {Max got up and had breakfast} rather than {Max had breakfast and got up} -- is in fact M-representational. The point is this: what is seen by relevance theorists as "linguistically encoded" is actually M-represented, but what is M-represented may be more than what is "linguistically encoded". Or rather, the distinction between what is M-represented and what is not M-represented may, but -- crucially -- need not necessarily, correspond to the relevance-theoretic distinction between what is "linguistically encoded/ decoded" and what is pragmatically implied/ inferred.

So it is not just the "linguistically encoded" concepts which are M-represented, but also some of the concepts which are assumed by Carston (1998) to result from pragmatic enrichment or loosening -- especially those which, according to Carston (ibid), contribute

to the explicit content of utterances. One of Carston's examples of enrichment (ibid: 309) is

(23) He wears rabbit.

Carston (ibid: 309) claims that {rabbit} in (23) "linguistically encodes" something like *rabbit stuff*. This concept *rabbit stuff* is enriched via strengthening by the hearer to *rabbit skin/ fur*. In response to this, it could be pointed out that there is no real conceptual necessity to suppose that {rabbit} "encodes" (i.e. conventionally-(2) M-represents) *rabbit stuff*. The notion that {rabbit} encodes something as general, uninformative and "uninteresting" as *rabbit stuff* is of course consistent with Sperber and Wilson's (1986a/ 95) argument (supported and developed in Carston 1998) that, because coded communication is subservient to inferential communication, "linguistic codes" need not and do not encode a good deal of what humans are interested in communicating. (See also my section 3.1.1.) But this means that the nature of Sperber and Wilson's "linguistic codes" is (at least partly) determined by the *non-linguistic* matter of what can be inferentially communicated. Another way of putting this is that relevance-theoretic "linguistic codes" seem designed to anticipate inferential processes. Furthermore, if "linguistic codes" are *for* coded communication, and subservient to inferential communication, then all the non-linguistic (and especially sociocultural) factors which bear on human communication in general must also bear on those so-called "linguistic codes". Within the relevance framework, it is on non-linguistic grounds -- i.e. the fact that anything more informative than *rabbit stuff* can be implied/ inferred -- that {rabbit} is assumed to "encode" the very uninformative *rabbit stuff*. What, then, is there to prevent {rabbit} from "encoding" the more informative *rabbit animal* for the similarly non-linguistic reason that someone simply intends {rabbit} to encode *rabbit animal*, in preference to having to imply/ infer it?

In fact it is *rabbit animal*, rather than *rabbit stuff*, which appears to be what the hearer/ reader of

(24) The rabbit ate the carrot.

would "automatically" recover from {rabbit}. So, if to "linguistically encode" is to "conventionally-(2) M-represent" in RC terms, I would argue that {rabbit} in (24) "linguistically encodes"/ conventionally-(2) M-represents *rabbit animal*. And I would argue that {rabbit} in (23) M-represents *rabbit skin/ fur* -- possibly non-iconically and non-conventionally, in the sense that the person who uttered (23) intended his hearer/ reader to calculate (i.e. pragmatically infer) *rabbit skin/ fur* from the conventional-(2) M-representatum *rabbit animal*.

However, it is possible that {rabbit} in (23) in fact "linguistically encodes" -- i.e. conventionally-(2) M-represents -- *rabbit fur/ skin*. One argument for this is that the hearer who hears/ reads {rabbit} immediately after {wears} seems more likely to "automatically" recover *rabbit skin/ fur*. There is also an independent reason for assuming that (23)'s {rabbit} "linguistically encodes"/ conventionally-(2) M-represents *rabbit fur/ skin*. Whereas the English {rabbit} can be used to refer to both animal and skin, the German and Putonghua equivalents can only be used to refer to the animal. This suggests that the relation between the English {rabbit} and the linguistic expression *rabbit fur/ skin* is conventional-(2), CSPR-specific.

Much of what I have just written about (23) also applies to some of Carston's examples of loosening (1998: 311). According to Carston, {heart} in

(25) Have you eaten my chocolate heart?

encodes something like *an anatomical heart*, from whose encyclopaedic entry the hearer picks out the properties which go with *chocolate*. An RC account of (25) goes like this. {Heart} does M-represent *an anatomical heart* in

(26) He has a weak heart.

but there is nothing to prevent it from M-representing (conventionally-(2) or otherwise) something like *heart-shaped thing* or *heart-like object* in (25). In fact, it seems more plausible for the second M-representation in constructions like {chocolate heart} -- e.g. {toy train}, {plastic elephant} -- to conventionally-(2) M-represent *train/ elephant/ X-like thing*, than for the hearer/ reader to have to loosen the relevant noun in each case.[8]

Similarly, it is possible that {bald} in Carston's example

(27) I love bald men.

conventionally-(2) M-represents a gradable concept rather than *totally hairless* -- contrary to Carston's claim that *totally hairless* is linguistically encoded, and the gradable *almost totally hairless*, the result of pragmatic loosening. The use of {bald} to M-represent *almost totally hairless* is certainly more common and usual than the use of {bald} to M-represent *totally hairless*. As Carston herself points out (1998: 327), a lexical concept like *totally hairless* is virtually never used "literally". Moreover, the general gradability of adjectives seems to be a convention specified by the conventional-(2) CSPR English -- hence M-representational constructions like {very bald} and {quite bald}.

---

[8] There is an alternative RC account of examples like {chocolate heart}. {Chocolate heart} may be the elliptical form of the M-representation {chocolate M-representation of a heart}, in the same way that the M-representation {That is T.S. Eliot}, produced in reference to the painting by Wyndham Lewis, is the elliptical form of {That is an M-representation of T.S. Eliot}. See Burton-Roberts 1994: 209-10 fn 18.

6.3 More on "linguistic codes" and "literal meaning"

Before I proceed, here once again is the main argument of 5.3: that linguistic expressions, in the form of syntactico-semantic objects or structured concepts, *ARE* LOT formulae. As I pointed out at the start of this chapter (in section 6), this obviously has implications for

(i) the notion of E-physical U "linguistically encoding" M; and

(ii) the notion of M being (taken to be) the "literal meaning" of E-physical vehicle U.

On the representational conjecture (RC), U is an intrinsically meaningless E-physical object. And it follows from my 5.3 argument that M is constitutive of some linguistic expression. Thus I maintain that what an individual considers to be the "literal meaning" M of some E-instantiable phenomenon U is in fact constituted by the structured concept -- the *linguistic expression* constituted by that structured concept -- M-represented by U. More precisely, the "literal meaning" of the E-instantiable phenomenon is the sum of (i) the cognitive value of the E-instantiable phenomenon's (intended and/ or assumed) M-representation of some structured concept, and (ii) the cognitive value constituted by as much of the M-represented concept as can be consciously accessed. I also maintain that a "natural language" or "linguistic code" like English is actually a non-natural, conventional-(2) CSPR. It follows from this that what is "linguistically encoded" by U is in fact what is non-linguistically M-represented by U -- i.e. M, a structured concept constitutive of some linguistic expression.

The aim of this section is to use the different types of M-representation to account for certain aspects of the relevance-theoretic notions of "linguistic codes" and "literal

meaning". To begin with, I have claimed that "natural/ external/ particular languages" or "linguistic codes" are really non-natural, extralinguistic, socioculturally differentiated, conventional-(2) CSPRs. Such a CSPR is not only generally agreed upon by the members of the social group which it distinguishes, but is also something which the members of that community are constantly exposed to, and thereby conditioned to use. Thus an individual would be able to "automatically" recover what is "linguistically encoded" -- i.e. the linguistic expressions whose M-representation is constrained by the conventional-(2) CSPR specific to his community.

This conventional-(2) CSPR may include 2-o and 3-o M-representations. For example, the individual who has been conditioned to use the conventional-(2) CSPR English would "automatically" recover *die* from utterances of {die} as well as of {kick the bucket}, and *the ordinary or average man* from {the man in the street}. He might also (as I suggested in the previous section) "automatically" recover either *rabbit animal* or *rabbit skin/ fur* from utterances of {rabbit}, depending on the M-representational context. In other words, {rabbit} may conventionally-(2) M-represent two distinct linguistic expressions, as do so-called "ambiguous expressions" such as {flying planes} and {port}. That {rabbit} or {flying planes} M-represents more than one linguistic expression is a matter of conventionally sanctioned M-representational indeterminacy. See Burton-Roberts 1994 and 1999b, and my section 5.1.

The main advantage of my non-natural, non-linguistic RC account of "linguistic codes" is that it is able to provide explanations for the more problematic -- because obviously non-linguistic and contextually determined -- features of the so-called "*linguistic* code". I have more than once pointed out that "linguistic decoding" only *seems* "automatic", insofar as the "decoding" individual is *conditioned* -- by *external*, mainly sociocultural factors and circumstances -- to use a particular "code" (i.e. CSPR). Furthermore, what is "linguistically encoded" is far from invariable. Carston mentions "disparate encodings" (1998: 82 fn 2) -- for example, {livid} "linguistically encodes"

flushed for some and *pale* for others. Most users of the CSPR English follow the rule which states that {derangement} "encodes" *derangement*. For Mrs Malaprop, however, {derangement} "encodes" *arrangement*. Non-fictional examples include {beg the question} (discussed in 6.1), and {disinterested}, which appears to "encode" *uninterested* for some and *disinterested* for others. My point is that "disparate encodings" can be used in different contexts by different individuals, or even by the same individual. This means that the recovery of "linguistically encoded logical forms" may be "automatic" in some cases, but may necessitate the less than "automatic" process of inference in those cases where different, unfamiliar or unknown "linguistic codes" are being used.

That the "linguistic code" is neither linguistic nor a code, but a non-linguistic sociocultural conventional-(2) CSPR, would be far more consistent with the fact that there are disparate "encodings" and "codes", and that some are more conventional-(2) or unconventional than others. It would also explain why "linguistic decoding/ encoding" is not as "automatic", and not as absolutely distinct from pragmatic inference/ implication, as Sperber and Wilson think it is. In fact, being non-natural and non-linguistic, so-called "linguistic decoding/ encoding" is as much a matter of what Chomsky (1995a: 30) calls "human interests, intentions, goals and actions" as pragmatic inference/ implication. The only significant difference between what is "linguistically encoded" and what is pragmatically implied is that the former is (assumed to be) agreed upon or known in advance, while the latter is not.

The "literal meaning" of an utterance is regarded by Chomsky and Sperber and Wilson as somehow based on, yet distinct from, the "linguistic semantics". Whereas the "linguistic semantics" of the utterance is seen as belonging exclusively to the "I-language" or "linguistic code", the expressing and recovery of "literal meaning" is supposed to involve both the "I-language/ linguistic code" and the performance/ central systems. Sperber and Wilson's (1986a/ 95) way of ensuring that the "linguistic semantics" of an utterance is distinct from the "real semantics" of the "literal meaning" is to claim that the

derivation of "literal meaning" involves the central *inferential* processes. For example, the "literal meaning" of the utterance of

(20) Max got up and had breakfast.

is not just the logical form "linguistically encoded" by the utterance, but also includes pragmatically inferred material -- namely, the concept of the temporal relation.

In my view, this relevance-theoretic notion of "literal meaning" is an attempt to combine certain pre-theoretical assumptions with a particular theory of what constitutes the linguistic. On the one hand, the classing of the pragmatically inferred concept as "literal" content appears to be based on the instinct (of audience as well as producer of the utterance) that the utterance of (20) *has* conveyed the concept of the temporal relation in some undeniable, explicit way. On the other hand, the relevance-theoretic distinction between the pragmatically inferred concept and the "linguistically encoded" content of (20) appears to be based on the notion that the concept of the temporal relation does not correspond to -- is not "encoded by", or the "meaning of" -- any E-physical property that is generally considered to be "linguistic". In the case of (21), however, the concept of the temporal relation is assumed by the relevance-theorists to be "linguistically encoded", apparently on the grounds that it corresponds to the E-instantiable "linguistic expression" {then}.

How Sperber and Wilson's (1986a/ 95) relevance theory takes all of the above into account appears to be as follows. Sperber and Wilson define the explicit content of an utterance as an *inferential* development of the logical form "linguistically encoded" by the utterance (ibid: 182). They claim that the explicature of an utterance of (20) contains both the "linguistically encoded" logical form and the pragmatically inferred concept. So, on the one hand, Sperber and Wilson acknowledge that the concept has to be inferred in virtue of *not* being encoded by some E-physical "linguistic" property. On the other hand,

they also relate the concept to the E-physical "linguistic" properties by incorporating the concept with the logical form which *does* correspond to -- in virtue of being "linguistically encoded" by -- the E-physical "linguistic" properties.

In the previous section I suggested that the concept of the temporal relation does in fact correspond to (and is M-represented by) some E-physical property -- the E-physical order of the conventional-(2) M-representations {Max got up} and {[Max] had breakfast} -- and is thereby explicitly conveyed. In the case of

(23) He wears rabbit

the explicit nature of the supposedly inferentially derived concept *rabbit skin/ fur* is more obvious. *Rabbit skin/ fur* could be regarded as more obviously explicit -- in virtue of having a more conspicuous E-physical vehicle/ representans, {rabbit} -- than the concept which is M-represented less conspicuously by the E-physical order of the M-representations in (20).

To posit an M-representational relation between an E-instantiable object and its "literal meaning" is a way of explaining the fact that "literal meaning" is explicitly conveyed, yet is sometimes also intended to be calculable/ calculated. Note that, while "linguistic codes" like English are conventional-(2) CSPRs, the M-representational relation between "literal meaning" and their E-physical vehicles may be of any possible sort -- conventional-(2), unconventional, or non-iconically non-conventional. Thus the "literal meaning" of some E-physical vehicle may be generally agreed upon -- a conventionally-(2) M-represented linguistic expression, like *die* being M-represented by {die} or {kick the bucket} according to the conventional-(2) CSPR English. It may not be generally agreed upon, as in the case of the "literal meaning" of {M-representation} or Mrs Malaprop's {derangement}, but nevertheless intended. And it may have to be inferred, if it is constrained by an unfamiliar CSPR, and especially if it is non-iconically

and non-conventionally M-represented. This fact that "literal meaning" is explicit yet possibly intended to be inferred is particularly obvious in the case of non-iconic non-conventional M-representation. Take {Max got up and had breakfast}. On the one hand, the recovery of what is non-iconically and non-conventionally M-represented (i.e. the concept of the temporal order of the events) is intended to involve calculation. On the other hand, the non-conventional M-representatum has as concrete an E-physical M-representans (i.e. the spatial or temporal order of the E-physical M-representations) as any conventional-(1) M-representatum, and can therefore be regarded as specifically corresponding to, even "encoded" by, that particular E-physical property.

I shall end this section with some general remarks on how unconventional and non-iconic non-conventional M-representation bears upon "linguistic codes"/ conventional-(2) CSPRs. Unconventional or non-iconic non-conventional M-representation may be entirely ad hoc, e.g. when an individual deviates from his conventional-(2) CSPR English for the particular occasion and purpose of conversing with Mrs Malaprop, or uttering {Max got up and had breakfast} with the intention of non-conventionally M-representing *Max got up and then had breakfast*. However, even a single ad hoc instance of unconventional or non-conventional M-representation, occurring within the framework of a conventional-(2) system of M-representation, would entail a conversion from the conventional-(2) system to a different M-representational system, in the way that a single malaprop would convert a Davidsonian prior theory into a passing theory (see Davidson 1986: 443, and my section 4.4). In other words, the individual who engages in unconventional or non-conventional M-representation in effect creates another system of M-representation, even if that system is used only on a single, particular and possibly very brief occasion.

Note also that both unconventional and non-conventional M-representations may -- in time, and under certain circumstances -- be *conventionalised*. As I argued in 6.1, most conventional-(2) higher-order M-representations (e.g. {the man in the street}, {pay through the nose}) are formerly non-conventional M-representations which have become

part of some conventional-(2) CSPR. Similarly, M-representations such as {internet}, {crack [cocaine]} and {quango} have become conventional-(2) in recent years, their initial unconventionality diminishing as they began to be used more widely and frequently. Conversely, conventional-(2) M-representations may become less conventional-(2) -- the M-representation of *fine* or *fastidious* by English {nice} being a case in point. It follows from all this that so-called "linguistic/ semantic change" is neither linguistic nor semantic, but has to do with changes in sociocultural M-representational attitudes and behaviour -- i.e. with people using different E-physical objects (or the same E-physical object, e.g. English {nice}) to M-represent different concepts at different points in time. Furthermore, the "particular languages"/ "linguistic codes"/ conventional-(2) CSPRs themselves cannot actually be said to *change*. Strictly speaking, they can only be *replaced by other CSPRs.* As I pointed out above, the smallest deviation from a particular CSPR constitutes the use of some other system of M-representation -- one which is slightly but nevertheless absolutely distinct from the first CSPR. Thus I argue that the name of a "particular language" -- "English", "Italian", "Swahili" -- does not refer to a single conventional-(2) CSPR, but is used as a label for a series of conventional-(2) CSPRs, each being replaced by the next with every addition or loss of some M-representational convention.

6.4 The non-M-representation of linguistic expressions and the relation between the linguistic and the language of thought

If the "literal meanings" of E-physical utterances are structured concepts which are (intended to be) M-represented by those utterances, then the "*non-literal meaning*" of an utterance could be regarded as consisting of a concept which is *not* M-represented by that utterance, but intended to be accessed through it. Note that this RC account is of "non-literal meanings" as concepts intended to be *not M-represented*, rather than *not intended*

to be M-represented. I take it that the use of an M-representation to access a concept which it does not M-represent is as non-natural and intentional as the use of an M-representation to access a concept which it does M-represent (conventionally or otherwise). In fact, I am suggesting that the deliberate non-M-representation of the linguistic expression which one intends one's interlocutor to access is as much of an -- albeit non-conventional -- M-representational strategy as the M-representation of a linguistic expression.

In the context of this thesis, the main significance of the phenomenon of non-M-representation is that it raises a question about the relation/ distinction between the linguistic and the LOT. For me, the logical conclusion of the representational conjecture (RC) is that linguistic expressions and LOT formule are ontologically identical, and in no way distinguishable from one another. There are others who subscribe to the RC, but remain convinced that there is a natural distinction between the linguistic and the LOT. For these individuals, non-M-represented linguistic expressions (especially those which are highly complex) would appear to be the obvious candidates for *LOT formulae which are not linguistic.*

The obvious examples of non-conventional non-M-representation have to do with what Sperber and Wilson (1986a/ 95) regard as *implicatures.* Take Sperber and Wilson's "strong implicatures" -- i.e. the implicatures

> ... which must actually be supplied if the interpretation is to be consistent with the principle of relevance, and for which the speaker takes full responsibility. (1986a/ 95: 199)

There is no doubt that the linguistic expression which B's utterance is intended to M-represent does not constitute all the concepts which B intends to convey.

(28) A: Do you ever speak to Charles?

B: I never speak to plagiarists.

The main propositional significance of B's utterance is, presumably, the sum of the M-represented linguistic expression (*I never speak to plagiarists*), the strong implicature

(29) *Charles is a plagiarist.*

and the deduction from the M-represented linguistic expression and implicature, namely

(30) *I never speak to Charles.*[9]

The point is that neither (29) nor (30) are M-represented, although both are accessed via B's M-representation of *I never speak to plagiarists.*

This non-conventional strategy of accessing linguistic expressions, not through their M-representation, but through the M-representation of other linguistic expressions, also appears to underlie at least some of what Sperber and Wilson call "weak implicatures". In the case of a relatively simple metaphor like

(31) Bill is a bulldozer.

the weak implicatures are mainly constructed from linguistic expressions conceptually corresponding to something like *an aggressive and domineering person.* These linguistic expressions are derived from the conventional-(2) M-representatum (*a tractor for clearing ground*) of the M-representation {bulldozer}, and from the hearer's/ reader's encyclopaedic entry for the M-represented linguistic expression.

---

[9] For Sperber and Wilson, both (29) and (30) are "strong implicatures".

All these implicatures are structured concepts which are intended/ assumed to be conveyed but not M-represented. For example,

(29) *Charles is a plagiarist.*

may not be M-represented or truth-conditionally related to what is M-represented, but in itself is (I argue) nevertheless a perfectly truth-conditional linguistic syntactico-semantic object.

However, there are weaker implicatures -- such as those conveyed by Virginia Woolf's definition of a highbrow as

(32) "... a man or woman of thoroughbred intelligence who rides his mind at a gallop across country in pursuit of an idea" (cited in Davidson 1979)

-- which some may regard as too complex and private to be *linguistic expressions*.

There is also BR's 1999b account of quotational mention, which appears to provide an example of an LOT formula which is not linguistic. Very briefly, he argues that the M-representation {the boss} in

(33) "The boss" is a noun phrase.

is used to identify the linguistic expression which is being *mentioned*, i.e. *the boss*, rather than to express the concept which the linguistic expression constitutes. In other words, the linguistic expression *the boss* is M-represented separately from -- and is not linguistically, syntactico-semantically related to -- the linguistic expression *is a noun phrase*. So, although {the boss} M-represents the linguistic expression *the boss*, and {is a noun phrase} M-represents the linguistic expression *is a noun phrase*, (33) as a whole is

not an M-representation of a linguistic expression. The proposition conveyed by the use of (33), therefore, appears to be a NON-LINGUISTIC LOT FORMULA which is NOT M-REPRESENTED BY (33)

But the non-M-representation and apparent un-M-representability of certain concepts does not constitute an argument for non-linguistic LOT formulae, simply because the fact of M-representation or non-M-representation has no bearing on the nature of what is or isn't M-represented (and vice versa). There are obviously LOT formulae which certain individuals choose not to M-represent, e.g. (29). There are definitely LOT formulae which have never been (and may never be) accessed, consciously or unconsciously -- including some infinitely complex LOT formulae. There are, presumably, LOT formulae to which certain CSPRs do not assign M-representations. It is possible that (33) is used to access a concept which has no conventional means of being M-represented. But it is in the nature of M-representation that anything can be M-represented by anything else. Thus, in principle, all LOT formulae are M-representable.

In my view, that some LOT formulae are M-represented while others are not (or never) M-represented does not follow from the nature of M-representation or the nature of LOT formulae, but from *why and how humans M-represent LOT formulae*. It has to be remembered that LOT formulae are not just the I-cognitive M-representata of the E-physical M-representations, but also *consciously accessed* through those M-representations. The M-representation of the inherently I-cognitive is obviously different from, say, the use of a drawing of a square to M-represent a building on a map, or the use of the letter "a" to M-represent the sound [a]. In each of the latter cases, the person who sets up the M-representational relation has prior access to both E-physical M-representans and E-physical M-representatum. On the other hand, if LOT formulae are what I believe they are -- (at least partly) constitutive of, and therefore inaccessible to, the central processes which give rise to consciousness -- then a person would be unable to intentionally pick out a particular LOT formula, or to reflect upon a particular LOT

formula, without the mediation of M-representations. Prior to the acquisition of some CSPR, his access to LOT formulae would be both unconscious and non-volitional.

That E-physical M-representations are used to provide conscious access to what they M-represent appears to be one of the reasons why many of the M-representations used by humans tend not to be wholly arbitrary. Rather, an M-representans tends to be chosen for some property or characteristic which can serve as an indication, an M-representational "clue", as to the nature of the M-representatum. For example, the "clue" may consist in a resemblance to -- and hence the iconic M-representation of -- the E-physical phenomenon which the M-representatum is the concept of. This appears to be the case for onomatopoeic expressions, some hieroglyphs and an M-representation such as {Max got up and had breakfast}.

Alternatively, an M-representans may hint at the M-represented linguistic expression or LOT formula -- call it L_1 -- in virtue of also being the conventional M-representation of another linguistic expression or LOT formula L_2 which is conceptually related to L_1. What I mean is that the accessing of L_1, though possibly "automatic", would in a sense be reinforced by prior knowledge of, or access to, the other M-representatum L_2. In 6.1 I referred to this phenomenon as second-order M-representation (e.g. {to laugh in someone's face}, its first-order M-representatum *to direct one's laughter straight towards someone's face*, and its second-order M-representatum *to openly show scorn for someone*).

But the most obvious example of the non-arbitrariness of M-representation is complex M-representation. Complex M-representations are constructed from simple M-representations which "[represent] what [they represent] without (and not in virtue of) representing any of its properties" (Burton-Roberts 1994: 198).[10] While there is the sort

[10] For those who subscribe to the notion of "particular languages", this distinction between simple and complex M-representations would be a distinction between "lexical items" (or even "morphological units") and "phrases/ sentences".

of M-representational convention which applies to single, simple M-representations (specifying, for example, the arbitrary M-representation of *dog* by {dog}), there is also the more general sort which determines the construction of complex M-representations from simple M-representations (e.g. the convention which states that {The dog chased the cat} M-represents *The dog chased the cat*, {The cat chased the dog} M-represents *The cat chased the dog*, and so on). These complex M-representations are more perspicuous than simple M-representations because they M-represent some aspects of the M-representata's *composition*.

Leaving aside the LOT formula expressed by the use of (33), the main characteristic of a supposedly non-linguistic LOT formula which is un-M-represented but intended to be expressed/ accessed is its highly complex, infinitely ramified nature. Take, for example, the LOT formula which was intended to be expressed, but probably not M-represented by,

(34) The worlds revolve like ancient women
Gathering fuel in vacant lots.
(from T.S. Eliot's *Preludes*)

Its full significance is never completely grasped, possibly not even by Eliot himself. Different readers in different contexts may access only parts of it, and different parts at that. Perhaps they may discover previously unaccessed parts. (They may, of course, also arrive at something which was not intended to be accessed.) But I suggest that, no matter how variable and vague it seems to be, this LOT formula is nevertheless constituted by a particular, perfectly precise, albeit infinitely complex concept. Furthermore, if such a concept is not M-represented, it is not because it is in itself un-M-representable, but because someone (in the case of (34), Eliot) has chosen not to M-represent it. On the one hand, a non-arbitrary conventional-(2) M-representation of an infinitely complex and ramified LOT formula would be humanly impossible to produce/ perceive, in virtue of its

correspondingly infinite E-physical dimensions. On the other hand, a non-conventional or unconventional M-representation which is E-physically manageable would be too arbitrary, and not perspicuous enough, in respect of the M-representatum's composition. So instead of M-representing the LOT formula -- i.e. accessing it through *its M-representation* -- one may choose to access it through *the M-representation of another LOT formula.*

So my argument is this: it does *not* follow from their being un-M-represented that certain LOT formulae are "*non-linguistic LOT formulae*" as opposed to "*linguistic LOT formulae*" or "*linguistic expressions*" -- simply because it does not follow from their being un-M-represented that they are un-M-representable. In my view, any LOT formula is M-representable, conventionally or otherwise, although whether or not it is actually M-represented depends on whether or not an individual chooses to M-represent it -- which, in turn, depends on the individual's intentions, abilities (cognitive, perceptual and motor), beliefs about his audience's cognitive and perceptual abilities (if an audience is involved), and general knowledge of the possibilities and limits of M-representation as a means of accessing LOT formulae. Furthermore, while LOT formulae tend to be un-M-represented (and accessed via M-representations of other LOT formulae) in virtue of being infinitely complex, that does not necessarily mean that they are therefore less accessible or "graspable" than LOT formulae which are regarded as "linguistic expressions" in virtue of being (conventionally-(2)) M-represented. As I have pointed out more than once, conscious access to concepts via M-representation is by definition indirect and incomplete. It follows from this that the M-representata of conventional-(2) M-representations -- especially of the simple, arbitrary, less perspicuous M-representations -- may not be any more accessible than, say, the LOT formula which is intended to be accessed via (but not M-represented by) (34). Take English {beauty}, {city}, or even {thing}. It is possible that humans have as little (conscious) idea of the true nature and conceptual extent of the M-representatum of {thing}, as they have of Eliot's un-M-

represented LOT formula. Or rather, it is possible that humans have as much conscious access to the true nature and conceptual extent of the un-M-represented LOT formula as they have of the LOT formula *thing*.

Thus I conclude, not only that all linguistic expressions are LOT formulae (as I claimed in 5.3), but also that all LOT formulae are linguistic (as I have defined "linguistic"). There may be a sort of distinction between "linguistic expressions" and "LOT formulae" in (M-representational) practice. An RC account of this distinction would be that the LOT formulae which humans use the E-physical to M-represent tend to be regarded as "linguistic expressions", while the LOT formulae which humans cannot or will not M-represent (conventionally or otherwise) tend to be regarded as "LOT formulae which are not linguistic expressions".

6.4.1 M-representation and reference

In 5.3 I argued that:

> (i) To be conscious at the most basic level is to have an unconscious, innate, non-volitional, unmediated -- and therefore direct and complete -- access to the structured concepts which constitute the central processes.

So, in order to be conscious -- i.e. to register, rather than to simply be impinged upon by, the E-physical world -- an organism has to have (the capacity to have) LOT formulae in the form of structured concepts. I also argued that

> (ii) To be conscious of being conscious (i.e. to be conscious of having concepts) is to have a conscious, acquired, volitional, M-representational -- and therefore

> indirect and incomplete -- access to (some of) the structured concepts which constitute the central processes.

By this, I mean that the capacity to use the E-physical to M-represent the I-cognitive affords a sort of consciousness different from -- and apparently higher than -- the fundamental, non-volitional, unmediated (vehicleless) form of consciousness described in (i). It has to be pointed out that the capacity for M-representation, as I conceive of it, is part of the central processes. Thus, on my definition of the central processes as being at least partly constituted by LOT formulae, the capacity for the M-representation of LOT formulae is itself constituted by LOT formulae.

In fact I would go as far as to claim that LOT formulae in the form of structured concepts are constitutive of all the thoughts, meanings, intentions, beliefs, desires, etc which the individual/ agent could be said to have -- whether the LOT formulae are non-volitionally and directly accessed, or volitionally and M-representationally accessed. Furthermore, I am of the opinion that these LOT formulae are innate, and purely internalist in nature. The question is, how does this bear upon, or account for, the individual's encounters with objects, situations and properties of the external world?

My views on thought and semantics are probably precisely what Putnam and his fellow "semantic externalists" would disapprove of, on the grounds that they are based on an erroneous tendency to " treat cognition as a purely *individual* matter", and to "ignore the social dimension of cognition" and "the contribution of the environment" (Putnam 1975a: 271). "Causal theories of meaning" such as Kripke's (1972) account of proper names and Putnam's (1975a/ b) account of natural-kind and physical-magnitude terms, on the other hand, proceed from the idea that

> To have linguistic competence in connection with a term... one must... be in the right sort of relationship to certain distinguished

> situations (normally, though not necessarily, situations in which the *referent* of the term is present). (Putnam 1975b: 199)

The "distinguished situation" for the "expression" {electricity},[11] for example, consists in an "introducing event" during which "a *description* of electricity is given, and generally a *causal* description" (ibid: 200). A "causal description" is

> ... one which singles out electricity as *the* physical magnitude responsible for certain effects in a certain way. (ibid)

The individual who knows how to use the "expression" {electricity} is linked by a "causal chain" to other users of the "expression", and ultimately to the "introducing event". Putnam concurs (ibid: 203) with Kripke's (1972) claim that a particular member of the "collective" linked by a causal chain need not have had "contact" with the referent, or even have "any good idea" of it. What is important is that this individual is linked by the causal chain to someone who has had "contact" with the referent.

Thus it is the EXTENSION -- the actual object, situation or property of the external world, to which an "expression" refers to -- which for Putnam and Kripke is the primary element of the "meaning" of the "expression" (or the primary element of the content of a thought). However, as Bilgrami (1992: 236) puts it, Putnam "bifurcates the intentional aspect of the mind into two" by claiming that an "expression" also has an INTENSION. Putnam (1975b: 200-1) concedes that the individual who uses an "expression" to refer to, say, electricity, would also have an intensional *concept* of electricity. This includes what is "linguistically associated" with the "expression". In the case of {electricity}, Putnam suggests that this may be no more than the concept of electricity as a physical magnitude, and capable of flow or motion. The intension of the "expression" {electricity} may also

---

[11] On the RC, any E-instantiable "expression" or "term" is an M-representation -- hence the braces around {electricity}.

include ever-changing concepts associated with "causal descriptions" of the extension of the "expression" -- from Franklin's, to the most recent accounts involving electrons etc.

Putnam also concedes (1975a: 246) that the intension of an "expression" does have a role in the use of the "expression". Unlike Kripke, who argues that a person does not need to have true beliefs about X to use a proper name to refer to X, Putnam argues that the person has to have some beliefs about the bearer of the name which are "true or approximately true" (1975b: 203). For example, someone who uses the name "Quine" to refer to the logician must have the right "linguistic ideas" -- that "Quine" is not a name which is restricted to females, for a start (1975b: 201). However, the "semantic externalism" of Putnam's approach is such that the intension of an "expression" does not have as crucial a role in the use of the "expression" as the *extension* does. As Putnam claims (ibid: 196-7), different individuals may use the "expression" to refer to the same thing, and yet have different intensional concepts of that thing.

Putnam criticises internalist theories of thought and semantics for failing to take into account the social and deictic aspects of meaning (e.g. in 1975a: 245-6, 271). Conversely, internalist theorists could criticise the externalists for ignoring the sortal difference between *extension* and *meaning*. If the extension or referent of an "expression" is what Putnam claims it is, *an object/ situation/ property of the external world*, then there is obviously a problem of how exactly the extension of an "expression" contributes to the "meaning" of the "expression", whether the "meaning" is regarded as a mental object or an abstract entity. In my view, since it is the "meaning" which is the subject of the theory or investigation, it must be the extension which has to be, as it were, adapted to the form of the "meaning" and not vice versa. The extension in itself, being E-physical, cannot be directly involved in, or constitute a part of, the abstract or I-cognitive "meaning" of an "expression". Thus I argue that it can only be the *mental (internal) effects* of an individual's *perception/ cognition* of

(i) the E-physical extension,

(ii) the E-physical effects produced by the extension (e.g. sparks produced by electricity), and

(iii) the *non-natural* relation between extension and "expression" inaugurated during what Putnam calls the "introducing event"

that constitute the part played by the extension in the "meaning" of the "expression".

In other words: when a referent is said to be assigned to an "expression", what is contributed to the "meaning" of the "expression" is not (cannot be) the E-physical extension or referent itself, but certain *perceptual and cognitive effects* which the user of the expression conceptualises in the same way that the so-called intension of the "expression" is conceptualised. Putnam objects (1975a: 271) to the internalist tendency to "ignore the *world*, insofar as it consists of more than the individual's 'observations' ". My point is that it is precisely the individual's "observations", and his concepts of those "observations", which constitute the only form of interaction which the individual is able to have with the "world".

Putnam seems to distinguish sharply between *concepts* and *extensions*, on the grounds that one's concept of, say, fish (e.g. *aquatic organism with gills*) is not a necessary and sufficient condition for the use of the "expression" {fish} to refer (or not to refer) to something (e.g. an aquatic organism without gills). But under my RC account of utterances and literal meanings, the "expression" {fish} is an E-physical construction which may be used to M-REPRESENT the concept *aquatic organism with gills*, the concept *aquatic organism without gills*, or even the concept which a particular individual has of a particular organism.

Similarly, {he} is (I argue) simply an E-physical object which is used to M-represent the concept *third person, male*, or a particular concept which the user of the M-representation has of a particular male person (say, Bill). The latter is what I consider to

be the result of "reference assignment". It has to be stressed that I regard the second concept to be as much of an I-cognitive concept (LOT formula) as *third person, male*. Conversely, *third person, male* is as much of an LOT formula, with as real a semantics, as the concept which arose from reference assignment. Note also that it does not follow from the use of the same M-representation (e.g. {Bill}) by different individuals to refer to the same person, that all those individuals are accessing or expressing the same concept of the referent. Just because the same object is perceived/ cognised by different individuals does not mean that they end up perceiving/ cognising it in the same way. For example, an individual A who only knows Bill by sight may conceive of Bill solely in terms of some of Bill's physical features, whereas the concept which Bill's close friend B has of Bill would probably involve much more than that.

Furthermore, neither of the LOT formulae which constitute A's and B's respective concepts of Bill is identical to the LOT formula which {Bill} M-represents in accordance with the conventional-(2) CSPR which A and B have been conditioned to use. The conventional-(2) M-representatum of {Bill} seems to be something like *person bearing the name of {Bill}*. A and B are socioculturally conditioned to recover the latter from occurrences of {Bill}, but through that conventional-(2) representatum are also able to access their individual (and probably more complex) concepts of the particular individual Bill.

A more complex case involves an M-representation like {the man on the beach}, which appears to be the conventional-(2) M-representation of the relatively simple concept *some man situated on some beach*. However, consider the following examples:

(35)(a) The man on the beach is insane.

(b) The man on the beach needs help.

Let's say that there is a man on the beach who happens to be Bill, Max's acquaintance. In the case of (35)(a), Max uses {the man on the beach} to refer to the man on the beach, who he thinks is behaving oddly, without knowing that the man is Bill. In the case of (35)(b), Max uses {the man on the beach} to refer specifically to Bill, indicating to a passer-by that Bill is, say, having a heart attack. My account of these examples is that, in both cases, Max would be socioculturally conditioned to use {the man on the beach} to access the concept *some man situated on some beach*. Through this conventional-(2) M-representatum, however, he would also be able to access his own particular concept of the man he has observed behaving oddly on the beach (in the case of (35)(a)), or his own particular concept of his acquaintance Bill (in the case of (35)(b)). The (35)(a) concept could be regarded as a non-conventional second-order M-representatum of {the man of the beach}, if it happens to be logically related to (e.g. if it happens to entail) the conventional-(2) first-order M-representatum *some man situated on some beach*. But the concept accessed by Max in the case of (35)(b) may bear no relation whatsoever to the first-order M-representatum. Such a concept could be regarded either as an unconventional or non-conventional third-order M-representatum of {the man on the beach}, or as a concept which is accessed via, but not M-represented by, {the man on the beach}.

Returning to Putnam's criticism of the failure of internalist theories of "meaning" to take into account the "social dimension of cognition" and the "contribution of the environment", my response is this. Firstly, I argue that any social aspects of thought and "meaning" have to do with the use of non-linguistic M-representations and M-representational conventions to access LOT formulae. Furthermore, the fact that the social conventions are M-representational implies that they are not as social as they might appear to be -- it follows from the opacity of M-representation that there is no way of telling whether two or more people associate exactly the same LOT formula with a particular M-representation. Secondly, I suggest that it is not the external world which

has a bearing on concepts/ LOT formulae, but the latter which are necessary for an organism to be able to register the external world and the so-called "extensions" of "expressions".

BIBLIOGRAPHY

Bilgrami, A. 1992. "Can externalism be reconciled with self-knowledge?". *Philosophical topics* 20, 233-267.

Black, M. 1972/3. "Meaning and intention: an examination of Grice's views". *New Literary History* 4, 257-79.

-- 1979. "How metaphors work: a reply to Donald Davidson", in Sacks 1979.

Blakemore, D. 1987. *Semantic constraints on relevance*. Oxford: Blackwell.

Bloomfield, L. 1926. "A set of postulates for the science of language". *Language* 2, 153-64.

Boyd, R. 1979/ 93. "Metaphor and theory change: what is "metaphor" a metaphor for?", in Ortony 1979/93.

Bromberger, S. 1989. "Types and tokens in linguistics", in A. George (ed.) *Reflections on Chomsky*. Oxford: Blackwell. Reprinted in Bromberger 1992.

-- 1992. *On what we know we don't know*. Chicago: University of Chicago Press.

Burton-Roberts, N. 1993. "Use, mention and linguistic occurrence: a representational approach". Ms. Newcastle upon Tyne.

-- 1994. "Ambiguity, sentence and utterance: a representational approach". *Transactions of the Philological Society* 92:2, 179-212.

-- 1998. "Phonology in UG?". Paper given at the LAGB Spring 1998 meeting, Lancaster.

-- 1999a. "Language, linear precedence and parentheticals", in P. Collins and D. Lee (eds.) *The clause in English: in honour of Rodney Huddleston*. Amsterdam: John Benjamins.

-- 1999b. "Ambiguity, quotation and the Grelling paradox: a representational account". *Newcastle and Durham working papers in linguistics* 5.

-- 2000. "Where and what is phonology?", in N. Burton-Roberts, P. Carr and G. Docherty

(eds.) *Conceptual and empirical foundations of phonology*. Oxford: University Press. In preparation.

-- & Carr, P. 1999. "On speech and natural language". *Language sciences* 21, 371-406.

Carr, P. 2000. "Nativism, autonomous phonology and sociophonetic variation", in N. Burton-Roberts, P. Carr and G. Docherty (eds.) *Conceptual and empirical foundations of phonology*. Oxford: University Press. In preparation.

Carruthers, P. 1996. *Language, thought and consciousness*. Cambridge: University Press.

Carston, R. 1988. "Implicature, explicature and truth-theoretic semantics", in R. Kempson (ed.) *Mental representations*. Cambridge: University Press.

-- 1996. "Enrichment and loosening: complementary processes in deriving the proposition expressed". *UCL working papers in linguistics* 8, 61-88.

-- 1998. *Pragmatics and the explicit-implicit distinction*. PhD thesis. UCL.

Chapman, S. 1998. *Accent in context: the ontological status and communicative effects of utterance accent in English*. Bern: Peter Lang.

Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.

-- 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

-- 1971. *Problems of knowledge and freedom*. New York: Pantheon.

-- 1972. *Language and mind*. 2nd edition. New York: Harcourt Brace Jovanovich.

-- 1975. *The logical structure of linguistic theory*. New York: Plenum.

-- 1980. *Rules and representations*. New York: Columbia University Press.

-- 1986. *Knowledge of language*. New York: Praeger.

-- 1987. "Lecture 2: Conceptual foundations of the study of language", in Chomsky, N. 1987. *Language in a psychological setting*. Sophia Linguistica.

-- 1992a. "Language and interpretation", in J. Earman (ed.) *Inference, explanation and other philosophical frustrations*. Berkeley: University of California Press.

-- 1992b. "Explaining language use". *Philosophical topics* 20, 205-31.

-- 1995a. "Language and nature". *Mind* 104, 1-61.

-- 1995b. *The minimalist program*. Cambridge, MA: MIT Press.

-- 1998. "Some observations on economy in generative grammar", in P. Barbosa et al (eds.) *Is the best good enough?: optimality and competition in syntax*. Cambridge, MA: MIT Press.

Davidson, D. 1973. "Radical interpretation". *Dialectica* 27, 313-28. Reprinted in Davidson 1984.

-- 1975. "Thought and talk", in S. Guttenplan (ed.) *Mind and language*. Oxford: University Press. Reprinted in Davidson 1984.

-- 1979. "What metaphors mean", in Sacks 1979.

-- 1981. "Communication and convention", in M. Dascal (ed.) *Dialogue: an interdisciplinary approach*. Amsterdam: John Benjamins. Reprinted in Davidson 1984.

-- 1984. *Inquiries into truth and interpretation*. Oxford: Clarendon Press.

-- 1986. "A nice derangement of epitaphs", in Lepore 1986.

Davis, S. (ed.) 1991. *Pragmatics: a reader*. Oxford: University Press.

Dennett, D. 1969. *Content and consciousness*. New York: Humanities Press.

Donnellan, K. 1966. "Reference and definite descriptions". *The philosophical review* 75, 281-304. Reprinted in Davis 1991.

-- 1968. "Putting Humpty Dumpty together again". *The philosophical review* 77, 213.

Dummett, M. 1986. "A nice derangement of epitaphs: some comments on Davidson and Hacking", in Lepore 1986.

Fodor, J.A. 1975. *The language of thought*. Hassocks: Harvester Press.

-- 1983. *The modularity of mind*. Cambridge, MA: MIT Press.

-- 1985. "Some notes on what linguistics is about", in J.J. Katz (ed.) *The philosophy of linguistics*. Oxford: University Press.

-- 1987. "Why there still has to be a language of thought", in *Psychosemantics: the problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

Frege, G. 1952. "On sense and reference", in P. Geach and M. Black (eds. and trans.) *Translations from the philosophical writings of Gottlob Frege*. Oxford: Blackwell.

Gathercole, S. and A. Baddeley. 1993. *Working memory and language*. Lawrence Erlbaum Associates.

George, A. 1989. "How not to become confused about linguistics", in A. George (ed.) *Reflections on Chomsky*. Oxford: Blackwell.

Grice, H.P. 1967. *Logic and conversation*. William James lectures. Harvard University.

-- 1975. "Logic and conversation", in P. Cole and J. Morgan (eds.) *Syntax and semantics 3*. New York: Academic Press.

Hacking, I. 1986. "The parody of conversation", in Lepore 1986.

Kaplan, D. 1989. "Demonstratives", in J. Almog, J. Perry and H. Wettstein (eds.) *Themes from Kaplan*. Oxford: University Press.

Katz, J.J. 1977. *Propositional structure and illocutionary force: a study of the contribution of sentence meaning to speech acts*. Hassocks: Harvester Press.

-- 1981. *Language and other abstract objects*. Oxford: Blackwell.

Kempson, R. 1986. "Ambiguity and the semantics-pragmatics distinction", in C. Travis (ed.) *Meaning and interpretation*. Oxford: Blackwell.

Kripke, S. 1972. "Naming and necessity", in D. Davidson and G. Harman (eds.) *The semantics of natural language*. Dordrecht: Reidel.

Lakoff, G. 1987. *Women, fire and dangerous things*. Chicago: University of Chicago Press.

-- 1979/ 93. "The contemporary theory of metaphor", in Ortony 1979/93.

-- and M. Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press

Lepore, E. (ed.) 1986. *Truth and interpretation: perspectives on the philosophy of Donald Davidson*. Oxford: Blackwell.

Lewis, D. 1975. "Languages and language", in K. Gunderson (ed.) *Language, mind and knowledge*. Minneapolis: University of Minnesota Press.

MacKay, A. 1968. "Mr Donnellan and Humpty Dumpty on referring". *The Philosophical Review* 77, 197-202.

Ortony, A. 1993. "Metaphor, language and thought", in Ortony 1979/93.

-- (ed.) 1979/93. *Metaphor and thought*. 2nd edition. Cambridge: University Press.

Papafragou, A. 1995. "Metonymy and relevance". *UCL working papers in linguistics* 7, 141-75.

Peirce, C. 1933. *Collected writings vol 4*. Cambridge, MA: Harvard University Press.

Putnam, H. 1975a. "The meaning of 'meaning' ", in Putnam 1975c.

-- 1975b. "Explanation and reference", in Putnam 1975c.

-- 1975c. *Mind, language and reality: philosophical papers volume 2*. Cambridge: University Press.

-- 1993. "Replies". *Philosophical topics* 20, 347-408.

Quine, W.V. 1960. *Word and object*. Cambridge, MA: MIT Press.

-- 1979. "A postscript on metaphor", in Sacks 1979.

Recanati, F. 1987. "Contextual dependence and definite descriptions". *Proceedings of the Aristotelian Society* 87, 57-73.

-- 1993. *Direct reference: from language to thought*. Oxford: Blackwell.

-- 1995. "The alleged priority of literal interpretation". *Cognitive science* 19(2), 207-32.

Richards, I.A. 1936. *The philosophy of rhetoric*. Oxford: University Press.

Sacks, S. (ed.) 1979. *On metaphor*. Chicago: University of Chicago Press

Searle, J. 1969. *Speech acts: an essay in the philosophy of language*. Cambridge: University Press.

-- 1975. "Indirect speech acts", in P. Cole and J. Morgan (eds.) *Syntax and semantics 3*. New York: Academic Press.

Soames, S. 1989. "Semantics and semantic competence". *Philosophical perspectives* 3, 575-96.

Sperber, D. and D. Wilson. 1986a/ 95. *Relevance: communication and cognition*. 2nd

edition. Oxford: Blackwell.

-- 1986b. "Loose talk". *Proceedings of the Aristotelian Society* 86, 153-71. Reprinted in Davis 1991.

-- 1990. "Rhetoric and relevance", in J. Bender and D.E. Wellbery (eds.) *The ends of rhetoric*. Cambridge, MA: Stanford University Press.

Wettstein, H. 1979. "Indexical reference and propositional content". *Philosophical studies* 36, 91-100.

Wilson, D. and D. Sperber. 1981. "On Grice's theory of conversation", in P. Werth (ed.) *Conversation and discourse*. London: Croom Helm.

-- 1993. "Linguistic form and relevance". *Lingua* 90, 1-25.

Wittgenstein, L. 1953. *Philosophical Investigations*. New York: Macmillan