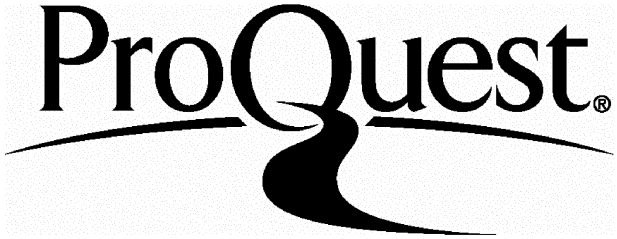
Pragmatics and the Showing-Saying Distinction

Timothy John Wharton

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

University of London, May 2003



Abstract

This thesis focuses on the role in communication of 'natural' signs and natural behaviours—facial expressions, spontaneous expression of emotion, interjections etc.—and considers how they might be accommodated within a pragmatic theory.

Linguists generally abstract away from such behaviours and focus on the grammar, or linguistic code. However, there are two reasons why the pragmatist should cast a broader net. Firstly, it is increasingly recognised in pragmatics that verbal communication is not simply a coding-decoding process but an intelligent activity involving the expression and recognition of intentions. Secondly, *sentences* are rarely uttered in a behavioural vacuum; the aim of a pragmatic theory is to explain how *utterances*—with all their linguistic and non-linguistic properties—are understood.

The analysis I propose has implications for theories of utterance interpretation. Firstly, verbal communication often involves a mixture of (what Paul Grice called) $\text{natural}_{(N)}$ and $\text{non-natural}_{(NN)}$ meaning, and there is a continuum of cases between 'showing' and meaning_{NN}; I show that natural communicative phenomena may be located at different points along the showing-meaning_{NN} continuum. This has clear implications for the domain of pragmatic principles or maxims, for it suggests that they are best seen as applying to the domain of intentional communication as a whole, rather than to the domain of meaning_{NN}. Secondly, the wider question of how 'natural' communicative behaviours are to be accommodated within a pragmatic theory breaks down into several further, more specific, questions—*How* are natural phenomena interpreted? Do they all work in the same way? I argue that some natural behaviours have a signalling function: they are, in effect, *natural codes*. Such behaviours do not fit easily into Grice's distinction between natural and non-natural meaning, which suggests that it is not exhaustive, and that it does not provide the best starting point for constructing theories of communication.

Pragmatics and the Showing-Saying Distinction

Table of Contents

Chapter Five

Chapter Six

Acknowledgments

I owe an enormous debt of gratitude to many people. I cannot hope to mention them all, and apologise to those I have left out. Thank you to my parents, Roy and Mary Wharton, and my brother Nick, for their love; Thank you to those few, close friends who have supported me in everything I have done—David Brown deserves special mention here, and Phil Harrison too, for happening to watch (and staying sober enough to remember) the Leonard Bernstein lecture that got him hooked on linguistics, and eventually led to my getting hooked too; Thank you to Katherine Bradley, for all her help in the past, and Peter Eastham for all his help in the present—special thanks also to Roma Loban, to whose memory this work is in part dedicated; Thank you to those teachers at school who recognised that education is (as Chomsky once said) 1% about methodology and 99% about inspiration—I single out Martin Herbert and Don Moffat here (though this is not to question their methodology); Thank you to everyone involved in the Department of Phonetics and Linguistics at University College London, where I have been studying and working for the past seven years—in particular I thank Neil Smith, whose book 'The Twitter Machine' reeled me in (having already been hooked). Thank you also to Nick Allott, Corinne Iten, Paula Rubio, Rosa Vega-Moreno and members of the Relevance Theory Reading Group for stimulating discussion, and Richard Horsey for sharing with me soft-shell crabs, crayfish popcorn and much discussion of cognitive linguistics in Santa Barbara, Summer 2001; Thank you to the UCL Graduate School who (somehow) deemed me worthy of investment, and provided me with the financial support that made it possible for me to continue my studies; Thank you to all the wonderful students I have taught at UCL; Thank you to Steve and Mary Skelly, who cut me

my own set of keys to their beautiful West London house, and to Clive and Denise Smith, who generously provided me with office space in Lewes.

I hope I will be forgiven if I also thank a man whose figure looms large over this thesis—a person I never met, but sometimes feel as if I have, and always feel as if I wish I had—Paul Grice.

As well as all the above, I would like to single out five special people and reserve my biggest thanks for them. I have been uncommonly lucky that these five people have found (and find) it within themselves to think I am worthy of their guidance, friendship and (in some cases) love. The first of these is my principal supervisor—Deirdre Wilson. To say she has been supportive over the past few years is to do no justice at all to everything she has done for me, both personally and professionally: on a personal level, she has helped me through some hard times, and shown patience, understanding and encouragement well beyond the call of duty; on a professional level, while some supervisors offer their students new roads down which to travel, Deirdre throws open the shutters to reveal whole new vistas, whole new landscapes to explore. I've really hardly begun my exploration, but I feel very lucky to have been given the chance.

The second person is Robyn Carston, my second supervisor. I single Robyn out because as well as all the time, effort and patience she has invested in me during my time at UCL as both an undergraduate and postgraduate student, Robyn is a teacher with the kind of qualities I allude to above. Studying her undergraduate courses was never work, it was always pleasure; one course in particular in the Spring of 1998 was largely responsible for my decision to make 'meaning' the general direction I intended to pursue.

And so to the final three: my wife Tamsin, and my daughters Xanthe and Zoë. I don't think I'm easy to be with at the best of times, and during the past few years (some of which have not been the best of times) I've been plain impossible. Tam, thank you for your love and support, and for making home more than just a place I live. Xanthe and Zoë, I'm *so* lucky to be your daddy; and among all the wonderful things that await you in your lives, I hope that one day you too can experience the sheer, indulgent pleasure of being allowed, no *encouraged*, to think and to learn. (Just don't leave it as late as I have.)

Introduction

> 'A wagging tongue ... proves to be only one part of a complex human act whose meaning must also be sought in the movement of the eyebrows and hand.'
>
> (Goffman 1964, pp. 133-134)

> 'We see it as a major challenge for any account of human communication to give a precise description and explanation of its vaguer effects. Distinguishing meaning from communication, accepting that something can be communicated without strictly speaking being *meant* by the communicator or the communicator's behaviour, is a first essential step. [...] Once this step is taken, we believe that this framework we propose... can rise to this challenge.'
>
> (Sperber and Wilson 1986/1995, pp. 57-58)

Sentences are rarely uttered in a behavioural vacuum. We colour and flavour our speech with a variety of 'natural' vocal and facial gestures, which indicate our internal state by conveying attitudes to the propositions we express or information about our emotions or feelings. Though we may be aware of them, such behaviours are often beyond our conscious control: they are involuntary, spontaneous. Almost always, however, understanding an utterance depends to some degree on their interpretation. Often, they show us more about a person's mental/physical state than the words they accompany; sometimes, they replace words rather than merely accompany them.

On the whole, the approach favoured by linguists is to abstract away from such behaviours, to sift out extraneous, paralinguistic or non-linguistic phenomena, and focus on the rule-based grammar—the *code* that constitutes language. This strategy has reaped rich rewards. Over the past thirty years linguists have suggested intriguing answers to the classical questions of language study (Chomsky 1986), and are now in a position to ask questions it was once not even possible to formulate (Chomsky 2000).

There are two reasons, however, why the pragmatist should cast a broader net. Firstly, thanks largely to the influential work of Paul Grice (1957, 1967, 1968, 1969, 1975, 1982, 1989),[1] it is now increasingly recognised that verbal communication is more than a simple coding-decoding process. Any attempt to characterise linguistic communication should reflect the fact that it is an intelligent, inferential activity involving the expression and recognition of *intentions*.[2] Secondly, the aim of a pragmatic theory is to explain how *utterances*—with all their linguistic and non-linguistic properties—are understood. The task, therefore, of describing and explaining precisely what is conveyed by natural communicative phenomena[3] such as those mentioned above would appear to fall squarely within the domain of pragmatics.

---

[1] Page references to these works are from Grice's 1989 anthology *Studies in the Way of Words* (hereinafter referred to as *Studies*). Where I think it might be of interest or relevance—in particular in Chapter Two—I correlate page numbers from *Studies* with the source of the quote in Grice's original work, whether that be the original typescripts of Grice's *William James Lectures* or the original published versions of those lectures.

[2] When I refer to human communication as 'intentional', I use the word in what Hauser (1996, p. 23) calls the 'rich, philosophical sense'. Thus, human communication is intentional insofar as it exploits the cognitive ability to attribute mental states—in particular intentions—to others. I point this out because later in the thesis I discuss various ethological concepts, and ethologists traditionally use 'intentional' in a different sense. '*Intention* movements' in non-human animals, for example, are movements that reliably predict a certain course of action: the term implies no pre-meditation on the part of the animal, and nothing 'intentional' in the 'rich, philosophical sense' on their part. Confusingly, it is the latter sense of the word that is adopted by *cognitive* ethologists (Allen and Bekoff 1997), concerned as they are with mapping the abilities of non-human animals to attribute mental states to others. (Although the notion of intentionality Allen and Bekoff prefer is that of Millikan 1984, whose view of intentionality—and human communication—is very different to the one presented in Grice 1957 and the one taken in this thesis.)

[3] I will switch to the term 'phenomena' from here on; this is because I have natural 'states' in mind—where a communicator is *being* something—as well as natural behaviours—where a communicator is *doing* something. See Allen and Bekoff 1997, pp. 41-48 for discussion.

Of the various issues raised in the forthcoming chapters, the question of how 'natural' phenomena are to be accommodated within pragmatic theory is a central one, and provides what I hope will be a discernible thread throughout the thesis. This wider question devolves into a number of further, more specific, questions—What is the relationship between natural communicative phenomena and intentional communication? *How* are natural phenomena interpreted? *What* precisely do they indicate?—each of which will be addressed at various points.

The kind of answers we provide to these questions will depend on many factors; more than most, however, they will depend on how we choose to characterise notions that are fundamental to the questions themselves: notions such as *natural*, *language*, *pragmatics* and *communication*. Before providing an overview of the structure of the thesis, then, a few remarks concerning these seem appropriate. The following are not to be regarded as definitions; they are simply nods in the general direction I intend to proceed in, and to that end I hope they will be regarded as helpful.

Regarding the term natural, it should be clear from my opening paragraph that what I have in mind is to contrast 'natural' phenomena, on the one hand, with human language, on the other. An immediate problem presents itself: the human linguistic code, after all, is of course itself entirely 'natural' (hence, 'human natural language'). This observation is central to the view of language adopted throughout what follows: language is not 'learned', it 'grows' (Chomsky 1988, p. 134). As well as this, I would not want to deny that the most 'natural' response in a given communicative situation is more often than not a *linguistic* one; so just as language is natural, so is language *use*. Perhaps more worryingly, there is of course a sense in which—as Mary Catherine Bateson (1996, p. 10) puts it—

'everything is natural; if it weren't, it wouldn't be. That's *How Things Are*: natural'.[4]

Strangely enough, the fact that nothing is *un*natural is part of the key to understanding the notion of 'natural' I intend. When I speak of 'natural', I am talking of phenomena that *mean* naturally, in the sense of Grice (1957): the antonym, then, is not *un*natural, but *non*-natural. Here 'means naturally' is roughly synonymous with 'naturally indicates', so in the same way that black clouds might be said to mean rain or spots mean measles, Lily's smile might be said to mean she is happy, or Jack's frown mean he is displeased. This can be clearly contrasted with the kind of meaning inherent in language (often referred to as arbitrary or conventional), which Grice called non-natural; so the word 'pluie' means 'rain'; 'Lily esta feliz' means 'Lily is happy', or what that remark meant was 'Jack is displeased'.

I should also point out that I will be focussing very much on a subset of phenomena that mean naturally. I will be mostly concerned with the kind of communicative behaviours or states alluded to in my opening paragraph: facial expressions, affective tone of voice, spontaneous expressions of emotion. In the subset I also include 'natural' gesticulation and manual gesture, with the (important) caveat that some gestures used in verbal communication are not natural in the sense I intend—the British two-fingered insult is a case in point—and with the further caveat that the kind of 'gestures' put to use by signers as part of the various deaf *sign-languages* are not—in a crucial sense—gestures at all:

[4] Dan Sperber (p.c.) once wrote to me in a similar vein that 'everything that is—or at least everything that is in time and space—is natural, including all things cultural, artificial, etc.'

they are part of language, and would fall on what Grice called the non-natural side of meaning.

It might be suggested that the above discussion could have been avoided had I chosen to adopt—rather than 'natural'—the terms 'paralinguistic' and 'non-linguistic', familiar from so much of the literature. I'm not convinced. For one, there is disagreement over what these terms mean. There are those, for example, for whom 'paralanguage' is only those *vocal* aspects of language use that are not strictly speaking language: intonation, stress, affective tone of voice, rate of speech, hesitation (if that can be considered to be vocal) etc. On this construal, facial expression and gesture are non-linguistic phenomena. Then there are others for whom the paralinguistic lines up with just about all those aspects of linguistic communication that are not language *per se*, but, are nonetheless somehow involved with the message or meaning a communicator conveys. People in the latter camp, presumably, have little use for the term 'non-linguistic' (or if they do, mean by it non-*communicative*, which confuses the matter even more, implying as it does that 'language' and 'communication' are the same thing). On the first construal, notice that while the set of paralinguistic phenomena intersects with the set of natural phenomena as construed here, there exist both paralinguistic phenomena that are not natural—*deliberate* frowns or *fake* smiles—and natural phenomena which might be co-opted for communicative use that I would not want to call paralinguistic on any conception—a bruise or a pale complexion, for example. In many ways, the second construal makes more sense to me; rising pitch is so often linked with rising eyebrows, for example, that I'm not sure why we would want to say that while the former is part of a paralanguage, the latter is not. I will, however, stick with 'natural'.

A final reason for preferring the term 'natural' is that it allows me (to a large extent at least) to abstract away from the debate on whether intonation and stress, for example, are part of language.[5] (Though I will later attempt to answer the question of whether or not interjections are part of language.) This is a complex debate—as far as I can see, one that is pretty much entirely dependent on how we define 'language'.

Oddly enough, many of those who use (and define) the terms 'paralinguistic' and 'non-linguistic' seem strangely content not to define *language*. They assume, I suppose, that it is easily definable, or that since we all have an intuitive handle on what 'language' is, a definition is not really called for. It's not clear to me that this is the case at all; if you're going to bandy terms like 'language' around, then you owe it to your reader to at least have a stab at telling them what it means to you. In what follows, then, I adopt what seems to me to be one of the clearer conceptions of language around at the moment: a resolutely Chomskyan one. I remain (relatively) undeterred by the knowledge that Chomsky himself may well regard aspects of the kind of enterprise on which I am embarking as a fruitless one (2000, pp. 19-74). However, I remain hopeful that if, as he once remarked, 'It is possible that natural language has only syntax and pragmatics' (2000, p. 132), then some of the following discussion might be of some value in clarifying the latter.

The view, then, is of language from a cognitive perspective. Language is an *I*nternal, *I*ndividual, *I*ntensional object—Chomsky's *I*-language. Humans have a dedicated mental 'organ' or 'faculty of language' (2000, p. 168)—potentially a *module* (or set of modules). In a normal person this will mature, given exposure

[5] See Gussenhoven (2002) for interesting proposals concerning this question.

to the appropriate environment, from an initial genetically-determined state, to a 'steady state' that can be said to represent knowledge of language. To be in this state is to know a certain set of rules or principles: language is a rule-governed (or principle-governed) system. It is also a creative, combinatorial system with a finite number of elements (morphemes), which can be combined to create novel utterances of arbitrary length. The set of rules or principles a speaker of a language knows constitutes a mental *grammar*, a *code* pairing phonological and semantic representations of sentences. Such a conception, of course, has its detractors; among these are some great thinkers—John Searle and Hilary Putnam to name but two. But it will not be my aim in this thesis to defend the Chomskyan approach (for discussion of some of the general objections raised see Chomsky 2000).

Turning to *pragmatics*, predictably enough my aim is to adopt a pragmatic theory that will provide an account of verbal communication—i.e. language *use*—that will complement the kind of internalist or cognitive approach to language adopted by Chomsky. On the question of precisely *how* this interaction between Chomskyan I-language and the cognitive construal of pragmatics I will outline below might take place, I will also not have much to say (see Carston 2002, pp. 1-14, for an overview of how a cognitive theory of pragmatics might fit in with broadly Chomskyan distinctions). There are various complications: not least that if, as the Chomsky quote above suggests, knowledge of language incorporates a pragmatic component, and if an explanatory linguistic theory is going to be—to use Chomsky's terminology again—a theory of a speaker's *competence* as opposed to *performance*, then it might be argued that our pragmatic theory should be a theory of pragmatic competence too. However,

while it is easy to see (even if you disagree with him on other issues) how an internalist approach to language ties in with the Chomskyan notion of competence, it is not entirely clear to me at least that a theory of language *use* should (or, indeed, can) abstract away from factors involved in a competent speaker's *performance* in the same way.[6]

Notwithstanding these issues, the pragmatic theory I adopt in this thesis—*Relevance Theory* (Sperber and Wilson 1986/1995)—does indeed take its lead from Chomskyan (and Fodorian) insights into language and mind, and is in the same spirit as their work. Relevance theory combines Gricean intention-based pragmatics with aspects of modern psychological research and cognitive science to provide a cognitive-inferential pragmatic framework, which is an abstract model of a communicator's *performance.* It takes as its domain a carefully defined sub-set of those cases that might—in folk terminology at least—be referred to as instances of *communication.* 'Communication' is a broad notion. As 'zoosemiotician' Walter Sebeok remarks:

> 'all organic alliances presuppose a measure of communication: Protozoa interchange signals; an aggregate of cells becomes an organism by virtue of the fact that the component cells can influence one another.'
>
> (1972, p. 39)

[6] Chomsky does not regard this as a problem; in an online interview (1999) he remarks: 'If we are using the term "competence" in my technical sense, then pragmatics is not part of a theory of linguistic competence, for uninteresting terminological reasons. If we are using the term "competence" in its ordinary English sense, then I suppose one might say that pragmatics is part of linguistic competence, but the conclusion is... uninteresting, merely a matter of terminology'.

Construed in this way, our pragmatic theory would indeed have to be what Chomsky (1992) has termed a 'theory of everything'; it would be required to encompass every possible facet of human interaction that might conceivably be said to be (in Sebeok's terms) 'communicative'; from socio-cultural right down to sub-personal phenomena: from fashion to pheromones.

The kind of communicative behaviour relevance theory concerns itself with is *ostensive* behaviour—behaviour by which a communicator provides evidence that they *intend* to communicate something. Language is seen as governed by a code. But utterance interpretation is a two-stage process. The linguistically encoded logical form, which represents the output of the mental grammar, is simply a starting point for rich inferential processes guided by the expectation that speakers conform to certain standards or expectations; that in (highly) intuitive terms, an audience knows that a communicator has a good reason for providing the stimulus which attracts attention to their intention to communicate, and that that reason is a good enough one for an audience to attend to it.

Fodor's chief objection to pragmatic programmes (an objection shared by Chomsky) is that the processes involved in utterance interpretation are processes of the 'global' kind; such processes are not amenable to scientific study (1983). In many ways, then, as Robyn Carston points out: 'the relevance-theoretic framework ... can be seen as a response to the challenge presented by these sceptics' (2002, p. 2). Indeed, rather than being under-pinned by the kind of global processes Fodor and Chomsky describe as best characterised by, for example, scientific theorizing, relevance theory is under-pinned by 'fast and frugal heuristics' of the kind currently gaining much currency in cognitive science. The proposal that humans are equipped to search for relevance does not

entail all-seeing, all-knowing, 'unbounded' rationality. Rather, evolution has left us with economical rules-of-thumb that enable us to make the most of our finite cognitive capacity. Furthermore, as we have seen, relevance theory has a carefully delimited domain; it is not even a theory of *communication* (let alone a 'theory of everything'). Throughout the thesis, I will provide examples of the theory in action, and hopefully go some way to showing that the framework Sperber and Wilson propose can rise to the 'challenge' described in the epigraph to this introduction.

In order to take the first 'essential step' referred to in that epigraph, I begin in Chapter One with a discussion of Grice's seminal paper—*Meaning* (1957). This is arguably one of the most influential papers of the past fifty years, and certainly one that has had a profound influence on linguists and members of the cognitive science community as well as philosophers.[7]

In this paper, Grice attempted to characterise non-natural meaning (meaning$_{NN}$) in terms of the expression and recognition of intentions; for Grice 'what is meant$_{NN}$' was broadly coextensive with what is intentionally communicated. I look at the distinction that Grice drew between the two types of meaning and present some of the tests he designed for recognising the distinction. These tests will be important to the discussion later on in the thesis. I then present an outline of Grice's original analysis and look more closely at the kind of intentions Grice regarded as central to his notion of meaning$_{NN}$. Crucial to this was the line Grice drew between 'showing' and 'meaning$_{NN}$' (or 'telling' or

[7] During his presentation at the *Mind and Language* Workshop in Oxford in September 2000, Alan Leslie remarked that it was Grice's 'Meaning' that was largely responsible for awakening his

'saying that'). I will argue that while intuitions may be divided on whether cases of 'showing' do or do not always amount to cases of meaning$_{NN}$, there seems little doubt that cases of 'showing' do indeed qualify cases of intentional communication of the kind a pragmatic theory should be considering: as Sperber and Wilson say in the epigraph, meaning and communicating do not always line up.

I then discuss some of the well-known problems with Grice's analysis. In particular, I focus on a problem arising with regard to the analysis of 'natural' behaviours, which—at least partially—stems from Grice's desire to distinguish between showing and meaning$_{NN}$. In Grice (1957) such behaviours were largely set aside (understandably so, given his aims). I argue that while it is clear that some natural phenomena do indeed *betray* our thoughts and feelings to others in a way we would not want to describe as intentionally communicating them (or 'communicating' them in *any* sense), some *are* used intentionally: they may not be deliberately produced, but that does not mean they cannot be deliberately (or intentionally) shown. Intentional verbal communication, then, involves *a mixture of natural and non-natural meaning.*

In Chapter Two I stay with Grice's work, and look in more detail at his notions of 'saying' and 'what is said'. This involves examining some of the intricacies (and history) involved in any attempt to provide an adequate characterisation. As well as discussion based on original typescripts of Grice's famous *William James Lectures*, it will also involve reference to the work of Grice's contemporaries at Oxford, including members of the 'ordinary language

---

interest in some of the psychological issues that have come to form the backbone of his research: '57 was indeed a good year—'Meaning', 'Syntactic Structures' *and* 'Jailhouse Rock'.

philosophy' movement, who along with Grice, were some of the pioneers of pragmatics. The discussion will involve looking at the distinction drawn by members of that school in the 1940s between 'describing' and 'indicating'. I also look in more detail at the cognitive implications of adopting an inferential model of communication based (loosely on Grice's insights). My conclusion is that the notions of 'saying' and 'what is said' most typically ascribed to Grice might not be the ones he had in mind.

In Chapter Three I present the central tenets of relevance theory, define *ostensive-inferential* communication, and clarify the notion of 'showing' in relevance-theoretic terms, with particular attention to how it relates to Gricean $\text{meaning}_{\text{NN}}$. I propose that there is a continuum of cases between showing and $\text{meaning}_{\text{NN}}$, and I argue that this has clear implications for what should be seen as the domain of pragmatic principles or maxims, for it suggests that they are best seen as applying to the domain of overt intentional communication as a whole, rather than to the domain of $\text{meaning}_{\text{NN}}$. Indeed, this is the approach taken by relevance theorists. The continuum (or at least various versions of it) will form a thread that will run through the thesis from this point on.

I further claim that the interpretation of many natural communicative phenomena is a vital component of understanding not only a communicator's attitude to the proposition they are expressing, but also the proposition that is being expressed. This interaction between the natural side of human communication and the explicit side, I will suggest, is one that has been overlooked and is in need of investigation. Finally, I introduce a distinction that will play a central role in the proposed analysis of natural behaviours: the distinction between translational and non-translational encoding, taking it's lead

from the distinction between *conceptual* and *procedural* encoding, first introduced by Diane Blakemore in 1987.

In Chapter Four I look at one possible way of exploiting both this distinction and the showing-meaning$_{NN}$ continuum developed in the previous chapter, and turn to partly natural communicative phenomenon: interjections—words (or, as the case may be, *non*-words) such as *ouch*, *wow*, *yuk* etc. These are felt by many to have a natural element to them in the sense that many seem to have developed as stylised exaggerations of entirely natural responses. It has been proposed, for example, that they occupy a position on a continuum between display and language proper (Goffman 1981). I consider this continuum in more detail, and argue that it might be seen as one version of the continuum outlined in the previous chapter. I discuss previous attempts to characterise the meaning inherent in interjections as conceptual, and show how that these are problematic. I therefore claim that the coded side of interjections is best dealt with utilising the notion of non-translational—or procedural—encoding, introduced in the previous chapter.

In Chapter Five I return to the kind of natural phenomena mentioned in my opening paragraph and suggest that a subset of these behaviours do not fit easily into Grice's distinction between natural and non-natural meaning. Many things in the world carry information, or 'indicate': tree-rings, footprints in snow, the scent of ripe fruit. We might include in these natural human phenomena such as shivers and smiles. However, only a sub-set of these indicators are ever exploited, and only a sub-set of *these* indicators have an indicating function, that is, owe their continued existence to the fact that they indicate. These distinctions hold for Gricean natural signs too. Some do not have an indicating function; I suggest that

the interpretation of these is governed entirely by inference. Others, however, do. I propose that these behaviours have a coded element and argue they are best analysed in terms of *natural codes*. The existence of these inherently communicative natural signs suggests that Grice's original distinction is not exhaustive. I then go on to suggest that the translational/non-translational distinction can shed light on the question of what kind of coding it is that natural codes employ. As was the case with the discussion of interjections in the previous chapter, this will also involve investigating alternative attempts at characterising natural codes.

The model of communication outlined in this thesis presupposes to a large extent that humans are endowed with considerable metapsychological prowess. In Chapter Six, then, I look in more detail at the role played by the attribution of mental states—or 'mindreading'—in human communication and interaction generally. I also return to the continuum outlined in Chapter Three, illustrated further in Chapter Four. I argue, in fact, that these are two separate continua. I explore the relationship between them, and suggest possible applications they might have: how *synchronic* versions might be seen as 'snapshots' of the kind of communicative stimuli humans have at their disposal, and how a *diachronic* perspective (on one of the two continua) might shed light on both historical linguistics, and also on the processes that underpin human communication from an evolutionary perspective. We end where we came in, with Paul Grice, and a 'mythical' account of the evolution of meaning$_{NN}$, which, I argue, may have implications for accounts of the evolution of language.

Chapter One

1. Natural and non-natural meaning

'Take Bach's *Well Temper'd Clavier*. To me it means molecular harmony. To my father, it means a broken sewing machine. To Bach, it meant money to pay the candlemaker.'

(David Mitchell—*number9dream*)

Among the ghosts that haunt the corridors of departments that profess (and foster) an interest in pragmatics, there are a great many philosophers of language. Though with the passing of time the influence of some of these has faded, there can be little doubt that the spirit of Paul Grice continues to exert a powerful influence. Not only was his work among the most influential in laying the foundations on which much of modern pragmatics is built, but his insights continue to provoke debate (and controversy). We may owe the term 'pragmatics' to Charles Morris (1938), but Grice certainly ranks highly among a select few to whom credit is due for shaping (and continuing to shape) the discipline as we know it today.[1]

To pragmatists, indeed linguists generally, Grice is remembered best for his *Theory of Conversation*, which he outlined in the *William James Lectures* delivered at Harvard University in 1967. But whilst this is a thesis with its roots

[1] Including, no doubt, John Austin, Peter Strawson and the later Wittgenstein. For a review of the history of the development of pragmatics as a discipline see Recanati (1998); for a review of the *pre*-history of pragmatics see the references therein.

firmly in pragmatics, it is another area of Grice's work—his *Theory of Meaning*, first outlined in 'Meaning' (1957)—that is the focus of this first chapter.[2]

Although the two 'theories' are often regarded as distinct, they are not unrelated. Indeed, it could be argued (and in the next chapter I will) that the two theories are mutually illuminating to the extent that we fail to do justice to either if we consider them independently of one another.[3]

It might also be noted that Grice's theory of conversation and his theory of meaning were but one part of a much larger programme; a programme Grice never finished (nor indeed could ever have hoped to). On the back cover of Grice's 1989 anthology *Studies in the Way of Words*, Simon Blackburn describes Grice as 'a miniaturist who changed the way other people paint big canvases'. I

---

[2] The paper was actually written some time before its publication. In his CUNY Ph.D thesis, Russell Dale (1996) remarks:

> 'Stephen Schiffer, Richard Grandy, and Richard Warner have all told me in personal correspondence that Grice originally wrote the paper for a seminar that he and Strawson were to give in 1948, but was reluctant to publish it. Strawson had the article typed out and submitted it for publication without Grice's knowledge. Strawson only told Grice after the article was accepted for publication. Stephen Schiffer has told me that Grice himself told him this story and Richard Warner has written to me that he also heard this story from Grice.'
>
> (1996: Chapter 1, *fn.* 31)

That Grice's paper—or certainly ideas that were central to the paper—was circulating at Oxford well before its publication is confirmed in a review of Holloway (1951) written by H. Hart in 1952 (see bibliography): in that piece, he remarks that 'in order that I should understand ...[a]...statement in the specific sense of "understand" appropriate to statements, it is sufficient (and necessary) that I recognise from the utterance what the speaker *intended* me to believe or do.' (p. 61) He then adds in a footnote (*fn.* 2) that it was a certain 'Mr. P. Grice' who first made this clear to him. Chapter 2 of Dale (1996) provides an excellent overview of the study of meaning in the twentieth century, from the work of Lady Victoria Welby (1893/1896/1911), to Ogden and Richards (1928) and Gardiner (1951), to Grice himself.

[3] I follow Stephen Neale here: 'It is at least arguable that the Theory of Conversation is a component of the Theory of Meaning. And even if this interpretation is resisted, it is undeniable that the theories are mutually informative and supportive, and that they are of more philosophical, linguistic, and historical interest if the temptation is resisted to discuss them in isolation from one another' (1992, p. 512).

respectfully disagree; while it's easy to form the impression that Grice was a miniaturist because of the capacity he had for taking infinite pains, I think that misses the point that, actually, the canvas he envisaged forming his work was—to coin a phrase used by Daniel Dennett—'Vanishingly Vast' (1995, p. 109).[4]

Grice's approach to meaning is neatly summed up by a quote from his 1989 *Retrospective Epilogue*: '...what words mean is a matter of what people mean by them.' (1989, p. 340). Meaning was to be understood in terms of propositional-attitude psychology; ultimately, the meaning of words reduced[5] to the beliefs, desires and intentions of communicators who uttered them. As well as shaping modern pragmatics, then, another part of Grice's legacy is one particular view of 'semantics', a view that has been pursued most notably in the work of Stephen Schiffer (1972)[6] and Russell Dale (1996). The view is summed up by Jerry Fodor as follows:

---

[4] I think this is one of the reasons that we should not be in the least surprised Grice's work is open to so many conflicting interpretations: we simply fill in the gaps that—given the breadth of his task—he was forced to leave, in a variety of different ways.

[5] Whether or not the term 'reduced' is appropriate to refer to the relationship between word meaning and propositional attitudes has been the subject of much debate. Schiffer takes it that Grice was indeed embarking on a reductionist programme; Avramides, on the other hand, argues it may be better to see Grice's analysis as a *reciprocal*, rather than a *reductive* one. Dale (1996) insists that Grice, although he did not necessarily support 'reductionism' *per se* (see 1989, p. 351), was nonetheless engaged in reductive analysis, and uses observations based on the fact that 'Meaning' was written nearly ten years before its publication (see *fn.* 2) as evidence to counter Avramides' claims, some of which rest on some comments made by Grice in the 1957 paper. Date of publication (as I have found to my cost) is often an unreliable source of data .

[6] Schiffer's 1972 book 'Meaning' originated as a doctoral thesis with Grice, and outlined an approach he later came to call *Intention Based Semantics*.

> 'English inherits its semantics from the contents of the beliefs, desires, intentions, and so forth that it's used to express, as per Grice and his followers... English *has no semantics*. Learning English isn't a theory about what its sentences mean, it's learning how to associate its sentences with the corresponding thoughts.'
>
> (Jerry Fodor 1998, p. 9)

'Meaning' was at least partly conceived as a response to Stevenson (1944), whose own theory of meaning, Grice argued, failed to capture the crucial difference between expression meaning and speaker meaning. Having failed to recognize this distinction, Stevenson's framework would never be able to capture what, for Grice was the very essence of meaning: that the linguistic meaning of expressions should ultimately be characterised in terms of speaker meaning—that words do indeed mean what people mean by them.

Grice was not the first to argue for this kind of view. Although we cannot be certain that he read her work (though it is appealing to think that he may have done), Grice was returning to issues that had first been raised over fifty years earlier by a certain Lady Victoria Welby:

> '...though we do now and then recoil from a glaring misuse of a term in the "rising generation" and lament such a lapse from *our* good ways, we never see that the fatal seed has been sown, that fatal tradition of a far more extensive misuse has been handed on, by us; that in scores and hundreds of instances we have carefully habituated the child, trained it, to say one thing when it means another...'[7]
>
> (1911, pp. 62-63)

[7] I'm not convinced that we *train* our children to say one thing and mean another (well, I'm convinced that we don't), but I hope the quote illustrates Welby's perspective on meaning and how it might have influenced Grice.

Grice begins his paper with an attempt to focus the audience directly on the specific type of meaning he was to be concerned with: *non-natural*$_{(NN)}$ meaning. In order to do this, Grice distinguished it from a notion he called *natural*$_{(N)}$ meaning. (Though he did also suggest that meaning$_{NN}$ might somehow be analysed in terms of meaning$_{N}$—a point to which I will return in later chapters.) The natural/non-natural distinction was intended to reflect what he later went on to describe as:

> 'a reasonably clear intuitive distinction between cases where the word "mean" has what we might think of as a natural sense, a sense in which what something means is closely related to the idea of what it is a natural sign for (as in "Black clouds mean rain"), and those where it has what I call a non-natural sense, as in such contexts as "His remark meant so-and-so".'
>
> (1989, p. 291)

Grice was not the first to consider this distinction. Effectively, he was carrying on a tradition that dates back through the work of Hobbes to at least as early as William of Ockham,[8] and arguably back as far as Plato's *physis-thesis* opposition, first characterised in the *Cratylus*.[9] In this dialogue, Socrates and Hermogenes discuss the origins of language, and debate whether word meanings are the result of some 'natural affinity' shared between form and meaning, or

---

[8] See *Summa Totius Logicae*, in which Ockham distinguishes examples of *naturaliter signifecare* from those of *signifecare per volunatariam institutionem*. He uses the following as examples of natural meaning: 'a stone means that wine is sold in the tavern', 'a smile means inner joy'. I may be wrong, but is not the first an example of *non*-natural meaning? My thanks go to J. L. Speranza (p.c.) for bringing Ockham's examples to my attention.

[9] See Hamilton and Cairns (1989).

whether they are simply conventionally agreed upon. (This too is a subject to which I will also return in later chapters.)

Although Grice's 1957 paper was primarily concerned with meaning$_{NN}$, it will suit my aims (for reasons that will emerge later on in the thesis) to dwell for a while on the notion of meaning$_{N}$. Before turning to Grice's characterization of the former, then, I will look at some of the tests Grice used to distinguish the two notions.

Grice began by providing examples of the two types of meaning. Consider (1) and (2) below:

(1) Those spots mean[10] measles.

(2) That remark means he has measles.

He then proposed a variety of ways in which the two types of meaning might be distinguished. In the first of these, he pointed out that cases of meaning$_{N}$ are *factive*. A person who says 'those black clouds mean rain' commits himself to the claim that it will rain (or has rained); in such cases, Grice says, *x means*$_{N}$ *p* or *x meant*$_{N}$ *that p* entail *p*. By contrast, cases of meaning$_{NN}$ are non-factive. A person who says 'his remark meant *it is raining*' does not commit himself to the claim that it is raining now, or, in fact, has been raining at all.

I originally found aspects of Grice's notion of factivity puzzling. For while it may indeed be that black clouds indicate a high probability of rain, it could be argued that they don't *entail* that it will rain. It might *not* rain, for example, or what those black clouds might actually mean is that there has been an explosion

[10] Despite the fact that I occasionally leave the audience implicit, 'mean' is a three-place predicate: the spots mean something to somebody.

at the oil refinery (or something). One way of understanding what Grice was getting at is to see the example sentences he uses as examples of utterances in which it is *true* to say that '*x* means *p*'. It is clear, for example, that when *U*'s utterance of 'Those black clouds mean rain' is true, then it will indeed always be the case that it rains (hence the entailment), since if it doesn't rain, then an audience might quite legitimately respond 'Well, it looks like those black clouds *didn't mean* rain after all'. The issue of speaker commitment runs parallel with this; so if *A* remarks to *B* 'That hissing sound means there is a snake under the table', she cannot reasonably add 'but don't worry, I don't believe there's a snake under the table', since by her use of the *natural* sense of 'mean' it is also entailed that she holds that belief. (This ties in exactly with Grice's comments 1989, p. 213.) Of course, what that hissing sound might *actually* mean is that there is a gas leak in the pipe immediately underneath the table, but that does not affect *A*'s commitment to the belief that there is a snake there, even though (parallel to the black clouds example), if there *isn't* a snake under the table, it is no longer *true* to say 'That hissing sound means that there is a snake under the table'.

Compare this with a scenario in which *U* asks *A* what *B* meant by the remark 'Il y a un serpent sous la table', to which *A* replies 'That remark means there is a snake under the table'. Here *A* can say this entirely independently of whether or not she believes that there is a snake under the table, and, regardless of whether there actually is a snake under the table or not, the remark will still *mean* (or have meant) the same thing (and *A* will still mean or have meant the same thing by her utterance of it). In cases of meaning$_{NN}$, then, a true utterance of '*x* means *p*' does not entail either the speaker's commitment to *p*, or *p* itself. This example uses a case of linguistic (or, as Grice called it, timeless) meaning, but I think that the

same point can be made with instances of *speaker* meaning, though I admit that here things become slightly less clear. For if *U* says to *A* 'You'd be mad to go out without your umbrella' (S_1) and by uttering S_1 *U* means that it's going to rain, there is a sense in which it could be said that *U* commits herself to that belief (that it's going to rain). This may be one of the reasons why I have seen people equate—I believe mistakenly—speaker meaning with *natural* meaning: Robyn Carston (p.c.) has confirmed to me that she has also come across this view; but it *cannot* be the case that they are co-extensive: firstly, *A*'s utterance only $means_{NN}$ what it does in virtue of the *intentions* behind the utterance; secondly, and more importantly, even if it doesn't rain, *A*'s remark *still* meant that it was going to rain (or, rather, *A* meant that it was going to rain by her utterance of S_1). As I mentioned above, if *A* utters 'Those black clouds mean rain', and it *doesn't* rain, then those black clouds *didn't mean rain after all.*

The notion of factivity was reflected in one of a series of tests, in which Grice contrasted ways in which utterances containing uses of the word 'mean' (in both senses) might be satisfactorily paraphrased. While (3a) is a plausible paraphrase of utterance (1), (3b) is not a plausible paraphrase of utterance (2). It may be true, but it does not convey the same sense of 'means' as that in the original utterance. (It may, in fact, paraphrase a parallel case of *natural* meaning.)[11]

[11] Of course, (3b) may also report an instance of speaker (as opposed to *expression*) meaning (a variety of $meaning_{NN}$). Notice again, however, that it does not follow from this that speaker meaning and $meaning_N$ are co-extensive. Consider: *U*'s remark only $means_{NN}$ that he_x has measles if *U* intends it to, and—in contrast to instances of $meaning_N$—*U* (or *U*'s remark) might still be said to have $meant_{NN}$ that he_x has measles even if, in fact, he_x doesn't have measles at all.

(3a) The fact that he has those spots means he has measles.

(3b) ??The fact that he made that remark means he has measles.

A further test that Grice proposed, this time for recognising $meaning_{NN}$, involved the paraphrasing of utterances such as (1) and (2) using direct quotation. (4a) below is not a satisfactory paraphrase of (1), but (4b) is a satisfactory paraphrase of (2).

(4a) ??Those spots mean 'he has measles'.

(4b) That remark means 'he has measles'.

In another test, Grice proposed that no conclusion about *what is (was) meant by (something)* could be drawn from an utterance that describes a case of $meaning_{N}$. Such a conclusion, however, can indeed be drawn from a case of $meaning_{NN}$ (see (5ab)):

(5a) ??What was meant by those spots was that he has measles.

(5b) What was meant by that remark was that he has measles.

Grice's primary concern, of course, was $meaning_{NN}$, and in particular, how the kind of meaning exemplified in (4b) and (5b) might be characterised in terms of intentions and the recognition of intentions. Later on in the lectures, Grice looked in more detail at various types of non-natural meaning. I look in more detail at these separate notions of $meaning_{NN}$ in Chapter Two.

Returning to the 1957 paper, Grice moved through a series of carefully constructed examples in order to identify precisely what type of intentions are required:

> 'A first shot would be to suggest that "*x* meant_{NN} something" would be true if *x* was intended by its utterer to induce a belief in some "audience" and that to say what the belief was would be to say what *x* meant_{NN}. This will not do. I might drop *B*'s handkerchief near the scene of a murder in order to induce the detective to believe that *B* is the murderer; but we should not want to say that the handkerchief (or my leaving it there) meant_{NN} anything or that I had meant_{NN} by leaving it that *B* was the murderer.'
>
> (1989, p. 217)

The problem in this case is that the 'communicator's' (or handkerchief-dropper's) intentions are entirely incidental to the 'audience's' (or detective's) response. The two are not linked in any way; nor can they be, since the 'audience' is entirely unaware of the 'communicator's' intentions. Grice then turns to a series of further examples, where—in contrast to the handkerchief-dropper example—the 'communicator' openly (hereinafter *overtly*) provides evidence of their intention to induce a belief:

> 'Clearly we must at least add that, for *x* to have meant_{NN} anything, not merely must it have been "uttered" with the intention of inducing a certain belief but also the utterer must have intended the "audience" to recognize the intention behind the utterance. [...]

> [(A)][12] Herod presents Salome with the head of St. John the Baptist on a charger.
> [(B)] Feeling faint, a child lets its mother see how pale it is (hoping that she may draw her own conclusions and help).
> [(C)] I leave the china my daughter has broken lying around for my wife to see.'
>
> (1989, p. 218)

For Grice, however, a problem remained. There is still a sense in the above examples in which the communicator's intentions are (at least partly) incidental to the intended response being induced in the audience. In (A), for example, Salome can infer that St. John the Baptist is dead solely on the strength of the evidence presented, and independent of any intentions Herod has in presenting her with his head (similar remarks carry over to (B) and (C)). Grice wanted to *distinguish* between merely (albeit overtly) drawing someone's attention to a particular object or a certain type of behaviour—'showing', which in his view did not amount to the object or behaviour meaning_{NN} anything (or anything being meant_{NN} by the 'show*er*'), and something being meant_{NN} by the object or behaviour in question (or by the person responsible for using the object or behaviour in a certain meaningful_{NN} manner):

> 'What we want to find is the difference between "deliberately and openly letting someone know" and "telling", and between "getting someone to think" and "telling".

[12] I've changed Grice's original (1), (2) and (3) here to (A), (B) and (C) respectively, to avoid confusion—the numerals (1) and (2) feature in the next quote.

> The way out is perhaps as follows. Compare the following two cases:
>
> (1) I show Mr. *X* a photograph of Mr. *Y* displaying undue familiarity to Mrs. *X*.
> (2) I draw a picture of Mr. *Y* behaving in this manner and show it to Mr. *X*.
>
> I find that I want to deny that in (1) the photograph (or my showing it to Mr. *X*) $meant_{NN}$ anything at all, while I want to assert that in (2) the picture (or my drawing and showing it) $meant_{NN}$ something (that Mr. *Y* had been unduly familiar), or at least that I had $meant_{NN}$ by it that Mr. *Y* had been unduly familiar. What is the difference between the two cases? Surely that in case (1) Mr. *X*'s recognition of my intention to make him believe that there is something between Mr. *Y* and Mrs. *X* is (more or less) irrelevant to the production of this effect by the photograph. Mr. *X* would be led by the photograph at least to suspect Mrs. *X* even if, instead of showing it to him, I had left it in his room by accident; and I (the photograph shower) would not be unaware of this. But it will make a difference to the effect of my picture on Mr. *X* whether or not he takes me to be intending to inform him (make him believe something) about Mrs. *X*, and not to be just doodling or trying to produce a work of art.'
>
> (1989, p. 218)

In any act carried out in which evidence is provided of an intention to 'induce a belief' or to 'inform', notice that there are *two* layers to be retrieved by the audience. The first, basic layer is the information being pointed out—in Grice's example, the fact that Mr. *Y* is indeed being unduly familiar with Mrs. *X*, and the second, the information that this first layer is being pointed out intentionally. In examples (A), (B) and (C) from the quote above, the communicator (Herod, the

child, Grice) provides overt evidence of their intention to inform (the second layer), but in these cases the basic layer of information is derivable without reference to this intention. For a case to count as one of $meaning_{NN}$ this basic layer should *not* be entirely derivable without reference to the second layer (and, furthermore, this should be intended). Grice concludes his formulation of $meaning_{NN}$ as follows:

> '"*A* meant something by *x*" is roughly equivalent to "*A* uttered *x* with the intention of inducing a belief by means of the recognition of this intention".'

This he later modified to the following (see Grice 1989, p. 92):

> '"*U* meant something by uttering *x*" is true iff, for some audience *A*, *U* uttered *x* intending:
> (1) *A* to produce a particular response *r*
> (2) *A* to think (recognize) that *U* intends (1)
> (3) *A* to fulfil (1) on the basis of his fulfilment of (2).'

In Lectures V, VI and VII of the *William James Lectures* (published as Grice (1968) and (1969)) Grice considered criticisms of the first of the above three clauses, which led him to make further modifications. This involved changing clause (1) to "'*U* meant something by uttering *x*" is true iff, for some audience *A*, *U* uttered *x* intending: '*(1) A to think that U thinks that p*". The effect of this was to allow Grice to distinguish between utterances in which the intention is not to induce a belief, but rather to get the hearer to think that the speaker holds a particular belief, and utterances uttered with the *further* intention of getting the

hearer to come to hold a belief (on the strength of the hearer thinking that the speaker holds a certain belief). Grice refers to these are *exhibitive* and *protreptic* utterances. (The latter are therefore protreptic as well as being exhibitive.)

Stephen Neale remarks of this modification:

> 'One worry about the suggested revision is that it does not comport well with the commonly held view that the primary purpose of communication is the transfer of information about the world: on the revised account, the primary purpose seems to be the transfer of information about one's mental states.'[13]
>
> (1992, p. 549)

Although these changes do not affect the arguments to come in the next section, I will return to them in the final section of my final chapter, when I come to consider the evolution of communication and language.

[13] Here, Neale is echoing a point originally made in McDowell (1980).

2. Showing and meaning

Un petit d'un petit
S'étonne aux Halles
Un petit d'un petit
Ah! degrés te fallent
Indolent qui ne sort cesse
Indolent qui ne se mène
Qu'importe un petit d'un petit
Tout Gai de Reguennes
(H. K. Van Rooten 1968, Verse 1)[14]

Grice's formulation of $meaning_{NN}$ inspired (and continues to inspire) a great deal of discussion. On the one hand, philosophers (including Strawson 1964, Searle 1965, 1969, 1979, Schiffer 1972) constructed a range of complex counter-examples. Many of these counter-examples posit the presence of ever higher-levels of intentionality (which rapidly induce a kind of psychic vertigo). These, and some of the possible solutions to them, are neatly summarised in Avramides 1989.[15]

However, there is another way in which the above formulation might be challenged, and it is this way I would like to explore. Instead of focussing on intentions *over and above* the basic intentions proposed by Grice, it looks *within* the formulation itself, and in particular at clause (3) of the above reformulation: the central role Grice saw for the second layer—the intention to inform—in deriving the first.

Schiffer addresses this point in his 1972 book *Meaning*:

[14] As Van Rooten notes 'The most fascinating quality of these verses is found upon reading them aloud in [a] sonorous, measured classic style' (p. (ii)—foreword). The reader might like to try this for themselves—preferably to someone else.

[15] There is also a brief discussion of the solutions proposed to the problems raised for Grice's original account by some of the counter-examples in Chapter Two.

'[O]ne thing that might be said is that in presenting Salome with the head of St. John the Baptist, Herod might mean that St. John the Baptist was dead. This does not strike me as a wildly implausible thing to say. Consider an analogous case.

(3a) *A*: "Let's play squash."
S: Holds up bandaged leg.

Here, I think, one would say, intuitively, that by holding up his leg *S* meant that he could not play, or that he could not play because his leg was injured; yet it would seem that the only difference between (3) [(A)—the Herod, Salome and St. John the Baptist example—TW] and (3a) which is possibly relevant is that the "inference" *A* has to make in the "bandaged leg" example is slightly less direct than in the case of St. John the Baptist's head, although in both cases one could make the relevant inference without any assistance on the part of *S*.

Grice has objected to me that while we may say that (in (3a)) *S* meant he could not play squash by holding up his bandaged leg, he could not mean thereby that his leg is bandaged. But, in the first place, even this is not an objection to the point I am trying to make, which is that there is no relevant difference between (3) and (3a), so that if we may say that *S* meant that he could not play squash, then—by parity of reason—we may say that Herod meant that St. John the Baptist was dead (it was not suggested that Herod meant that there was a severed head on his charger). In the second place, I think that it is false that *S* could not mean that his leg was bandaged by holding up his bandaged leg. Consider (3b):

(3b) *A*: "I've heard that your leg is bandaged. Is it true?"
S: Holds up bandaged leg.

Here, I think, one would say that *S* meant that his leg was bandaged.'

(1972, p. 56)

Schiffer's argument, then, is that cases such as (1) from the above quote (the photograph example)—and, indeed, cases such as (A), (B) and (C) from the quote before that—should be regarded as instances of meaning$_{NN}$.

Francois Recanati also addresses this point, in his 1987 book 'Meaning and Force':

> 'Take Grice's example of Herod bringing to Salome the severed head of John the Baptist. By this "utterance", the "speaker" S (Herod) openly intends to provide A (Salome) with reason to believe that the following conditions obtain: John is dead, and S wants A to share this knowledge. Why should this not be considered a case of communication? Grice's reason for excluding this case is that for (an important part of) the speaker's intention to be fulfilled, it is not necessary that the intention be recognised: The severed head of John the Baptist, *by itself*, is evidence that he is dead, and to conclude that it is so, A does not have to recognise S's intention. Grice is right to point out that there are two sorts of cases: cases in which only the speaker's intention is intended to provide evidence (this is what Grice calls "non-natural meaning", and it is indeed central in linguistic communication) and cases in which the 'utterance' is intended to provide evidence over and above the evidence provided by the speaker's intention. But there is no reason, it seems to me, to restrict the label "Gricean communication" to the first category of cases, however important they are.'
>
> (1987, p. 189)[16]

My argument will run along similar lines, although as I mentioned in my introduction, my intention (echoing the point made by Recanati) is to try and

[16] See Neale 1992, pp. 547-549 for further discussion.

characterise intentional communication, not meaning$_{NN}$. While Schiffer and Recanati focus on the St. John the Baptist example (A), I will focus on example (B), which will lead me on to my discussion of natural behaviours and their role in intentional communication.

For a communicative act to be *intentional*, I will argue, the important thing is that evidence is provided of an intention to inform, and not whether in the absence of such an intention, an audience might have been able to draw their own conclusions. As I will show, the very fact that a communicator has provided evidence of an intention to inform may lead the audience to make 'less direct' inferences.

Consider (6a-e) below—adapted from Grice's example (B) above. In all these cases something has happened that has produced a response in an audience:

(6a) Mary is asleep. Her mother notices that she is pale and concludes she is unwell.

(6b) Feeling unwell, Mary lies in bed with her eyes closed. She intends her mother to see how pale she is but really doesn't care if this intention is noticed or not.

(6c) As (6b), except that here Mary's mother instinctively guesses at Mary's intention that her mother sees how pale she is.

(6d) Feeling unwell, Mary deliberately and openly lets her mother see how pale she is, so she will notice and help.[17]

[17] The original wording in Grice's example is 'Feeling faint, a child lets its mother see how pale it is (hoping that she may draw her own conclusions and help)'. I mention this because I would rather let the reader know 'deliberately and openly' that my 'adaptation' is precisely that. There are three key changes from the original, none of which, to my mind, affects the arguments

(6e) Mary says to her mother 'I don't feel well'.

As I have pointed out, Grice noticed that before we can be said to be dealing with a case of meaning$_{NN}$ certain intentions must be present. Firstly, the response itself must be intended—this rules out (6a) as a case of meaning$_{NN}$; secondly, the audience must recognise the intention to produce that response—this rules out (6b); thirdly, the communicator must intend that the audience should recognise the intention to produce that response—this rules out (6c). The final all-important condition, the one that rules out (6d), and makes (6e) a case of meaning$_{NN}$, is that only in this example does Mary intend that the recognition of her intention to produce the desired response will play a crucial role in producing the response itself. In (6d) Mary's mother *can see for herself* that Mary is unwell.

No one would propose that the scenario described in (6a) is a case of intentional communication. Mary is asleep; she does not intend to communicate anything. This might be better described as a case of *accidental information transmission*: Mary's pale complexion shows her mother that she is unwell. In fact, even to propose that this is communication is to use the word extremely

---

to come, but which simplify the point I am trying to make. The first change is the introduction of the phrase 'deliberately and openly'. Since he is contrasting this example with a previous one in which an 'utterer' is not overt about their intention to inform (the 'handkerchief-dropper' example), it seems clear that what Grice had in mind in his original example was a case in which the child (in my adaptation, Mary) intends to have her intention recognized. In fact, the phrase 'deliberately and openly' comes from Grice's own characterisation of the 'feeling faint' example (and others) in the next paragraph (1989, p. 218—and see below). The second change is the omission of the phrase 'draw her own conclusions'. I don't think that if the child is acting 'deliberately and openly' in showing her mother that she is pale, then the mother *is* 'drawing her own conclusions' any more (*cf.* cases of *accidental information transmission* mentioned below). The third change is the omission of the word 'hoping'. I made this change to avoid having to get into any discussion about whether 'hoping' necessarily involves 'intending'.

broadly. Intuitively, we would be loath to say that an individual walking down the street *communicates* every piece of information a passer-by might infer from his physical appearance, his demeanour, his clothes etc.

It is less obvious in (6b) and (6c) that we are not dealing with full-blown intentional communication. After all, Mary does intend to inform her mother of something. However, she is not being open about this informative intention, and while she might indeed be said to be communicating intentionally, she is certainly doing so covertly. (We might compare this with the 'handkerchief dropper' example.)

But what of the cases in (6d) and (6e)? While there is certainly a sense in which Mary's mother can see Mary's pale complexion and draw her own conclusions *irrespective* of Mary's intentions, I think (echoing Schiffer and Recanati) that there are good reasons to suggest that (6d) might be regarded as an instance of intentional communication (though—as Recanati suggests—this concept needs to be distinguished from meaning$_{NN}$).

Firstly, Mary is being 'deliberate and open' about her intentions. Even if she only intended to inform her mother that she was feeling unwell, by being 'deliberate and open' she is certainly being overt about her informative intention, rather than covert as in (6b) and (6c). There is a clear sense in which it is *Mary* showing her mother she is unwell, rather than just her pale complexion (as in (6a)). Secondly, and more importantly, notice that Mary does not just intend to inform her mother that she is unwell, but also that she wants her mother to help. If Mary's mother *does* in fact infer this, I think we would be loath to say that her inference is entirely down to her having drawn her own conclusions, and not, to some extent at least, the result of inferring intentions Mary had. For in general,

someone who is 'deliberately and openly' letting someone know something creates the expectation in their audience that they have done so for a reason: in order to have their informative intention fulfilled, a communicator must first let her audience know that she has such an intention in the first place. This reflects the point made by Schiffer in the quote above: the inference from 'Mary has a pale complexion' to 'Mary wants help' is *less* direct than the inference from 'Mary has a pale complexion' to 'Mary is unwell'. The *motivation* for making this less direct inference is the very fact that Mary—by acting deliberately and openly—has created an expectation in her mother that there is something extra to infer.

3. Deliberately shown natural behaviours

In my introduction I mentioned three questions, the first of which concerned the relationship between naturally occurring behaviours—natural signs in Grice's sense—and intentional communication. The above discussion suggests one way[18] in which we might explore the role natural behaviours play in intentional communication.

In the following, I am talking primarily of cases where an individual *does* something—cries, shivers, smiles—as opposed to *being* something—pale, or covered in spots. I will discuss three examples: the first of these I regard as fairly unproblematic; the second I regard as potentially problematic, but plausible nonetheless; the third is an example from a group of natural behaviours on which I intend to concentrate in Chapter Five. I would also add that I am not proposing

[18] There are other ways, one of which is the focus of Chapter Five.

it is always the case that natural behaviours are deliberately shown; just that the fact that they can be, and sometimes are, reveals an under-explored area in the study of human intentional communication.

Consider crying. Crying is a natural sign that someone is distressed or unhappy. Parallel to example (6a), it is not hard to imagine a case in which we see someone crying (perhaps in the street, or in a restaurant), and the information that has been transmitted to us, that this person is in distress or unhappy, has been transmitted (at least to us) entirely accidentally. Furthermore, I think we can also imagine (or recall) cases where, despite their best efforts to conceal them, tears[19] betray the true feelings of someone we are talking with. Perhaps they cover their face with their hands, or turn away, or hold something up in front of their face. Or perhaps they sit there crying, trying desperately, but failing, to hold back the tears.

However, I think it is equally true to say that we can all imagine or recall cases where there has there been *no attempt* by the person we are talking with to hold back the tears: cases in which a person is, in a sense, crying openly. In doing so (parallel to (6d)), a person might intend to inform their audience of their distress, and by displaying their natural behaviour their informative intention is

[19] I am abstracting away from the issue of the degree to which the 'tears' I speak of as being hidden, held back, or openly shown are themselves a natural sign of distress, and how precisely they are linked with crying. After all, we can shed tears without 'crying' (to expel a foreign body) and cry without shedding 'tears' (the crying of infants is a case in point—for a few weeks/months a crying baby sheds no tears). Once young children have learned that tearful crying is a successful strategy for gaining the attention of their parent or carer, tears appear to take on a (slightly sinister) manipulative function. So often, if a child hurts herself, she will only start tearful crying when her parent or carer arrives at the scene.

made evident to both communicator and audience.[20] The fact that a communicator has done so creates in their audience the expectation that there is something extra to infer: perhaps, for example, the *reason* this person is crying.

It is important to stress at this point that I am talking of behaviours that are *shown* deliberately, and not those that are *produced* deliberately: involuntary, spontaneous behaviours that are voluntarily shown, as opposed to voluntary behaviours. In the above discussion I have been very careful to refer to crying 'openly' as opposed to crying '*deliberately* and openly' for this very reason. The phrase 'crying *deliberately* and openly' brings voluntary behaviours to mind in a way that is potentially confusing. I am not concerned here with *faked* natural behaviours[21] or the tears of actors.[22]

Grice himself was perfectly happy to regard the voluntary production of otherwise natural behaviours as meaning *non*-naturally:

> 'For consider now, say, frowning. If I frown spontaneously, in the ordinary course of events, someone looking at me may well treat the frown as a natural sign of displeasure. But if I frown deliberately (to convey) my displeasure, an onlooker may be expected, provided he

[20] I say '*might* intend' because I would not like to deny that there are probably cases in which crying openly does not amount to an intentionally (deliberately) shown act—expression of grief at a funeral, for example.

[21] Another issue worthy of consideration is that actually there are two ways in which a facial expression, for example, might not be wholly spontaneous: that is, it may be deliberately *produced* or it may be deliberately *concealed*. These, together with further problems that arise with a simple spontaneous/voluntary—deliberate/involuntary dichotomy such as the one I am presenting, are discussed in Ekman 1997. A further consideration is that what on a given occasion is perceived as deliberate showing might actually be the result of a communicator *suppressing the concealment* of a certain behaviour.

[22] Though this last is an interesting one: many actors, asked to play a scene in which they are expected to cry, rerun emotional crises in their mind in order to recreate *real* tears.

> recognizes my intention, *still* to conclude that I am displeased. Ought we then not to say, since it could not be expected to make any difference to the onlooker's reaction whether he regards my frown as spontaneous or as intended to be informative, that my frown (deliberate) does *not* $mean_{NN}$ anything? I think this difficulty can be met; for though in general a deliberate frown may have the same effect (with respect to inducing belief in my displeasure) as a spontaneous frown, it can be expected to have the same effect only *provided* the audience takes it as intended to convey displeasure. That is, if we take away the recognition of the intention, leaving the other circumstances (including the recognition of the frown as deliberate), the belief-producing tendency of the frown must be regarded as being impaired or destroyed.'
>
> (1989, p. 219)

Grice is surely right on this. The deliberate production of an otherwise natural behaviour (in order to make evident an intention to induce a response or belief) is a clear example of intentional communication (and for Grice, $meaning_{NN}$). But it is natural, involuntary, spontaneous behaviours that are the focus here, and *their* status with regard to intentional communication that is being considered; just because a behaviour may not have been deliberately produced, the argument goes, it does not follow that it cannot be deliberately (or intentionally) *shown*.[23]

[23] Incidentally, I should add that it also does not necessarily follow that, as one interpretation of the above quote from Grice would suggest, a deliberate frown necessarily communicates displeasure: is there not such a thing as an ironic frown? Also, an argument parallel to the one Grice gives in the above quote could not strictly speaking be applied to spontaneous and deliberate *smiles*, for example. As I discuss briefly below, the spontaneous and the deliberate smile involve different muscles (and different neurology is implicated), and are hence quite easily perceived as 'real' or 'faked'. It would be interesting to investigate whether the same kind of differences exist between real and fake frowns. (I don't know of any such study.)

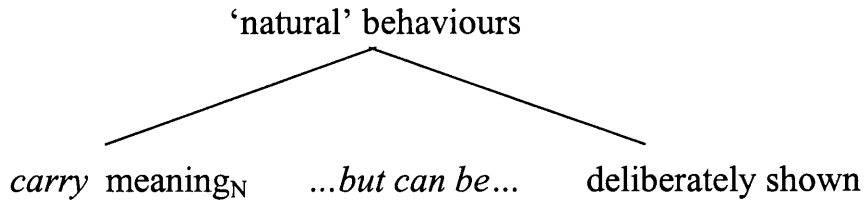
The second example is a shiver. Jack and Lily are sitting outside a London café on a typical bright spring day in London. It's freezing cold, Lily is miserable, and she wants to go inside. She feels herself beginning to shiver, looks at Jack, and draws his attention to her involuntary shiver. As with the deliberately shown crying, an expectation is created in the audience, for in providing evidence of her informative intention, Lily draws attention not only to the fact that she is cold, but also to the fact that she is *cold enough to want to go inside.* Again, the inference from Lily's natural (but deliberately shown) shiver, which indicates 'I am cold', to 'I want to go inside' is, to adopt Schiffer's terminology once more, 'less direct'. I find this example plausible too, but it is potentially problematic. For shivers are extremely transient things. There is therefore a great deal of potential for a shiver to be exaggerated, developed and stylised to the point where we might characterise it as being deliberately *produced* as well as deliberately shown. In which case, as we have seen, there is no doubt that the shiverer can be said to be communicating intentionally. (In Chapter Four I propose that interjections such as *ouch* and *wow* might have originated and developed as stylised exaggerations of 'natural' vocal gestures.)

The third example is a smile, which together with facial expressions in general will be the focus of the Chapter Five. If we accept that crying and shivering can be shown to an audience deliberately and openly, it is only a short step to accepting the same of involuntary, spontaneous smiles (and other natural facial expressions). Jack gives Lily a bunch of flowers, and Lily responds by letting Jack see her spontaneous reaction—a smile. Of course, smiles are also susceptible to being exaggerated, developed or faked. In the case of smiling, though, the task for the audience of detecting such differences is made easier by

the fact that spontaneous and deliberate smiles differ in various ways. Evidence from clinical neurology suggests that the neural pathways involved in spontaneous and deliberate facial expressions are different (Rinn 1984), and physiological evidence supports the claim first made by Duchenne (1862/1990), and later Darwin (1872/1998), that the muscles used in spontaneous smiles are different to those responsible for deliberate ones. Ekman (1992) remains unconvinced as to whether or not false facial expressions are always detectable to an audience, but in the case of the smile they surely are: firstly, fake smiles always mean *non*-naturally (see Grice's earlier remarks about a behaviour's 'belief-producing tendency'), and an audience must infer the intentions behind them—ask yourself when the last time was that you asked yourself 'what *did* he mean by that entirely natural smile?' (probably never), in contrast to the last time you asked yourself 'what did he mean by that phoney smile?' (probably not long ago); secondly, fake smiles stick out like sore thumbs—take a look at the photos of yourself at that *awful* wedding you went to a few years back.

I conclude, then, that behaviours that might, from a Gricean viewpoint, be regarded as simply cases of natural meaning, can also be deliberately shown, and hence used in intentional communication. This is not to suggest that they do not convey information in other ways (i.e. accidentally). However, it does enable us to make a clear distinction between those cases alluded to in the introductory chapter, where natural behaviours *betray* our mental/emotional state, and those in which these behaviours are recruited for use in intentional communication. These observations might be represented as in *fig. 1* below:

fig. 1



I will later propose that some natural behaviours (which Grice only regarded as carrying meaning$_N$) have an indicating (more precisely, *signalling*) function. This suggests that his natural/non-natural distinction is not exhaustive, and it follows that the right-hand side of the above diagram will require some fine-tuning.

In later chapters I will claim that there is a continuum of cases between showing, construed as above, and 'saying' (later to be modified to 'meaning$_{NN}$). At the showing extreme of this continuum lie clear cases of spontaneous, natural display (such as those mentioned above); at the other extreme lie clear cases of 'saying', where the evidence provided by a speaker is a linguistically encoded form. In between, lie a range of cases in which natural, conventional and coded behaviour mix to various degrees (e.g. pointing, stylised expressions of emotion, interjections). In the next chapter I turn to the 'saying' end of the continuum, with reference once more to Grice's work.

Chapter Two

1. Saying and Meaning

1.0 The territory

No formal training in linguistics is required to know that whatever 'saying' is, it does not necessarily amount to the same thing as 'meaning'. Most people share the intuition that what a speaker says may fall short of what they mean, or that by (or in) saying one thing a speaker might mean another thing entirely. Of course, some people always say exactly what they mean (to the awe-struck admiration or horrified embarrassment of people—such as myself—who often find it difficult to say what they mean), but this is not trivially true, and the fact that it is not is further intuitive evidence that saying and meaning are distinct.

But whilst they can be shown to be distinct, what a speaker says and what a speaker means are, of course, interrelated: there is a sense of 'saying' on which you can't say anything without meaning something; and if you didn't mean anything, it's hard to see how you can be regarded as having said something. So although saying (and the derivative notion 'what is said') will be the main focus of this chapter, it will be necessary at times to underpin the discussion by looking at some of the ways in which an utterer or an utterance might be said to mean something. In fact, one of the main premises on which the arguments in this chapter rest is that those aspects of Grice's work that deal with 'saying' and those

that deal with 'meaning' are mutually illuminating; it might even be argued that his work on saying forms *part of* his work on meaning.[1]

In any thesis which adopts a continuum between 'showing' and 'saying that' as one of its central threads, some discussion of what precisely constitutes 'saying' is required. Since this issue is also an aspect of Grice's work that regularly attracts discussion (not to say controversy), my aim in this chapter is to discuss two questions:[2]

(A) What did Grice mean by 'say'?

(B) What did Grice mean by 'what is said'?

These are important questions; saying was central to Grice's work, particularly to that area of his work concerned with distinguishing what speakers say from what they imply or *implicate*: without some concept of saying his famous distinction could not have been drawn at all.

Despite the centrality of the notion to his work, Grice is often regarded as having failed to provide an adequate characterisation of saying. Matters are not helped by the fact that in his attempts to do so, he appears to be drawn in two different directions. In the first, he admits that he is using (or at least seeking) a 'favored', to some degree 'artificial' (1989, p. 118) sense of *say*, and devotes time to developing a formal analysis. But as Stephen Neale points out, Grice only

[1] See *fn.* 3 of the previous chapter (p. 23); Grandy and Warner (1986, pp.1-45) also see Grice's work on 'saying' in the light of his wider philosophical programme, including his work on reason and rationality (e.g. Grice 2001, Chapters 1-4).

[2] During the writing of this chapter, it's become increasingly clear to me that even though the answer to Question (B) depends entirely on the answer to Question (A), it helps to consider them separately.

takes us so far with this analysis before reaching what might be construed as 'a rather disappointing terminus' (1992, p. 557). Certainly, what is widely regarded as his final analysis of saying—broadly speaking, those facets of speaker meaning that overlap with sentence meaning—leaves crucial issues unresolved. In the second direction, Grice is true to his 'ordinary language philosophy' roots, and admits he will 'have to assume to a considerable extent an intuitive understanding of the meaning of *say*' (1989, pp. 24-25). This, though, is equally problematic. For while it's easy enough to make intuitive distinctions between saying and meaning of the kind described in the introductory paragraph of this paper, it is by no means easy to give a precise and exhaustive *intuitive* characterization of saying on its own.[3] There's no small amount of irony in the fact that any attempt to characterise Grice's notions of saying and what is said is largely a matter of working out what he *meant* by what he said (and wrote) about it.

There is, at least, fairly broad agreement that Grice intended what is said to coincide with the truth-conditional content of an utterance, or the proposition expressed by a given utterance of a sentence on a particular occasion. Many theorists have claimed, however, that Grice's notion falls some way short of satisfying this intention. In a nutshell, the current thinking goes, not only do we often mean more than we say, but—if what is said is to be coextensive with

[3] Recanati (1993) argues that *what is said* is a consciously accessible level, on which users do have reliable intuitions. I agree with Robyn Carston that 'the variety of views that have been expressed in the literature [...] reflects the variability of intuitions' (2002, p. 168), and with Wilson and Sperber (2002. p. 620), who 'doubt that there is any common sense notion of *what is said* capable of playing a useful role in the study of verbal comprehension'. This doesn't mean, however, that we shouldn't make an effort to get straight exactly what Grice meant when he used the notion.

truth-conditional content—we typically say$_{(T\text{-}C\ CONTENT)}$ more than we say$_{(GRICE)}$, and what is said$_{(GRICE)}$ does not exhaust what is said$_{(T\text{-}C\ CONTENT)}$. As a result, a great deal of work since Grice has been devoted to bridging a perceived gap between his notion of what is said and the truth-conditional content of an utterance. Philosophers and linguists have supplied countless examples to demonstrate that such a gap exists, and various theoretical tools and corresponding terminological devices have been proposed in attempts to bridge it.

It is not my aim in the following discussion to question any of the advances that have been made in this area. Far from it, what follows can only lend support to them. Nor, incidentally, do I hope to rehabilitate the notions of saying and what is said. In exploring Grice's notion of saying, however, I would like to question the received wisdom on where Grice himself might have stood in the debate which his work has prompted. This will involve charting a grey area between the notions of saying and what is said commonly ascribed to Grice, and those which I will suggest are closer to the ones he had in mind. At the very least, I hope to convince the reader that there is evidence that Grice—were he still here to participate—would have gone along with many of the recent philosophical and linguistic attempts to bridge the gap between say$_{(GRICE)}$ and say$_{(T\text{-}C\ CONTENT)}$; at most, I hope to show that for Grice the gap simply did not exist. It may also be, as I will show at the end of this chapter, that he may have alluded to the fact that natural phenomena might contribute to 'what is said' and suggested a way in which they might have been accounted for in terms of his original analyses.

1.1 Saying and truth-conditions

Despite the fact that Grice avoided the term 'truth conditions', it is generally agreed that his notion of what is said was to have coincided with the notion of the proposition expressed by the speaker,[4] or the truth-conditional content of an

[4] Though Saul (2002) takes a different view. Saul argues that some post-Gricean enterprises (in particular, relevance theory) have misinterpreted Grice's notion of *what is said*. I too will suggest that there has been widespread misinterpretation of what Grice meant by what is said, but not for the reasons Saul suggests. Saul claims that the notion of what is said that interested Grice was a *normative* version—what is *actually* said—as opposed to what the speaker intended to say, or what the hearer takes the speaker to have said. But given that notions such as belief, desire and intention are so central (not to say fundamental) to Grice's work, whether on meaning, philosophical psychology or reason and rationality, it seems highly implausible to suggest Grice only had some abstract, speaker-and-hearer-independent notion in mind.

As well as failing to mesh with the overall fabric of Grice's philosophy, it's hard to see what theoretical significance a normative conception of what is said might have. It might conceivably be of some use in a Court of Law, where the innocence or guilt of the accused can turn on what was actually said in an utterance of an ambiguous sentence (*cf.* the famous British case of Derek Bentley and Christopher Craig, where Bentley, on seeing his friend drawing a gun on a policeman, uttered the ambiguous 'Let him have it!'—Craig fired the gun, fatally injuring the policeman). But notice that even in situations such as this, the cases for the prosecution and the defence will depend on what the hearer could be legitimately presumed to have taken the speaker to have intended to say; yes, this may coincide with, or underpin, some normative notion of *what is said*, but it still has a reference to the speaker's intention built in.

It has independently been suggested to me that a notion of what is *actually* said might be of relevance in cases of slips of the tongue etc. This is an interesting point, but is of course irrelevant to any discussion of what Grice meant by 'say', since for Grice saying entailed meaning, and in slips of the tongue nothing was actually said at all. As a further consideration, I wonder about cases in which a speaker is deliberately ambiguous and intends two different interpretations to two different hearers. Suppose a particularly demanding student—*X*—says to me 'Tim, I'm trying', and I reply to the student (within earshot of my teaching colleague, who also teaches *X*) 'Yes, we all know that you're trying'—intending the student to interpret me as saying 'We all know you—*X*—are making an effort', and my colleague to interpret me as saying 'We all know you—*X*—are very demanding': what is actually said here? Do I actually say two things? Do I only actually say one thing, and remain unaware of the other? Where some normative level of meaning is absolutely *vital* is at the level of semantics, but there is an increasing consensus that we cannot expect the semantics to deliver something co-extensive with *what is said* i.e. something fully propositional or truth-conditional (see Section 2 of this chapter).

utterance. The following quotes from Steve Levinson, Herb Clark and Stephen Neale (echoing my earlier remarks about Grice's failure to provide a definitive characterisation) are representative of this view:

> 'Grice uses the phrase *what is said* as a technical term for the truth-conditional content of an expression, which may in fact be somewhat less than the full conventional content.'
>
> (Levinson 1983, p. 97*n.*)

> 'What is said (in Grice's special sense) is what speakers mean mostly through the conventional content of the sentences they utter—indeed, through only that part that affects the truth of their utterances.'
>
> (Clark 1996, p. 141)

> 'Although Grice is not as explicit as he might have been, it is clear upon reflection (and from scattered remarks) that *what is said* is to do duty (with a proviso I will get to in a moment)[5] for *the statement made* or *the proposition expressed* by *U*. Where the sentence uttered is of the type conventionally associated with the speech act of asserting (i.e. when it is in the indicative mood) what is said will be straightforwardly *truth-conditional.*'[6]
>
> (Neale 1992, pp. 520-521)

The second part of Saul's paper, which—since it falls beyond the scope of this thesis—I will not address in any detail, is a sustained attack on relevance theory. It's true that if the version of relevance theory she attacks were the one proposed by Sperber and Wilson, then some of Saul's arguments might have some force. However, it isn't, so they don't.

[5] Neale's proviso is that for something to be a part of what *U* said, it must also be a part of what *U* meant.

[6] I think it only fair to point out that this quote is not representative of Neale's final position on 'saying' in his 1992 paper, which pays close attention to the points of contact between Grice's theories of Meaning and Conversation in a manner not typical of some commentators on Grice in the literature.

Thus, what is said is to be distinguished not only from what is implied—or conversationally implicated—but also from conventional *non*-truth conditional meaning. So in (1) below, the contrast introduced by the word 'but' between 'Xanthe is eight' and 'she's very tall' does not form part of the truth-conditional content of the utterance, and hence is not part of what is said:

(1) Xanthe is eight, but she's very tall.

According to Grice, words such as 'but' and 'moreover', and expressions such as 'on the other hand', *conventionally* implicate[7]—or indicate—the performance of a higher-order, '*non*-central' speech act (of, for example, 'contrasting').

This analysis is reminiscent of the speech-act distinction between describing and indicating (Austin 1962, Searle 1969, 1979); indeed, Grice's analyses of discourse connectives fit comfortably into the broader speech act framework. Under the speech act approach, sentences both express propositions, which describe the world (corresponding to truth-conditional content), and may contain non-truth-conditional indicators, which *indicate* the speech ('illocutionary') act a speaker is intending to perform, or the propositional attitude a speaker is intending to express. The difference in meaning between (2abc) below is captured by proposing that although all three sentences have the same propositional, or *descriptive* content— Zoë goes to nursery at time *t*—they differ

[7] A failure to adequately separate what is said from what is conventionally implicated is one of the criticisms Neale offers of Grice's final analysis of saying (1992, p. 555).

in their illocutionary force:[8] (2a) has the force of a question; (2b) of a request for action; (2c) of an assertion.

(2a) Does Zoë go to nursery?

(2b) Zoë, go to nursery!

(2c) Zoë goes to nursery.

Despite the similarities with aspects of the speech-act approach, and obvious affinities with speech-act philosophers, Grice had good reason to use the terminology he did, and hence avoid the term 'truth-conditions'. As Neale points out (1992, p. 556), '[Grice] cannot make a direct appeal to truth-conditions for fear of undermining one part of his project'. That project (or at least part of it) was to characterize sentence meaning as depending on a convention[9] among speakers to use certain words with certain *intentions*. To appeal directly to truth-conditions would have been inconsistent with such a view. In Grice's picture, truth-conditions are a derivative notion. Rather than being the starting point for an analysis of meaning, they are a property of utterances only (ultimately) in

[8] This distinction between propositional/descriptive content and illocutionary force is one version of John Austin's distinction between locutionary and illocutionary acts. See Recanati (1987, pp. 236-266) for detailed discussion of Austin's distinctions and the various interpretations of his work. This is not a debate I intend to get too involved in—though I will return to aspects of it in Section 2.3 of this chapter; nonetheless, the confusion and uncertainty that abound serve as a useful reminder that nothing should be taken for granted in the philosophy-of-language neck-of-the-woods (and thus form a neat backdrop to the (re)interpretive work being attempted in this chapter).

[9] Though Grice was not completely happy with the notion of convention. As Neale points out: 'Grice's use of the word 'conventional' in "conventional meaning" should not be taken too literally, for it is Grice's view that linguistic meaning is not to be explicated in terms of what other philosophers might think of as convention' (1991, p. 520 *n.*).

virtue of the intentions behind those utterances. As Grice himself puts it in the Retrospective Epilogue of *Studies*: 'what words mean is a matter of what people mean by them' (1989, p. 340).

Although it seems clear that Grice did intend what is said to equate with truth-conditional content, on one reading at least he appears to have had a fairly minimalist view of what saying (and hence what is said) actually was. As Herb Clark puts it:

> 'But what is saying? According to Grice—though he was vague on this point—it is the literal meaning of the sentence uttered with its ambiguities resolved and its referents specified.'[10]
>
> (Clark 1996, p. 143)

This widely held view can be traced back to an often-quoted section of Lecture II of the *William James Lectures* delivered at Harvard in 1967, and published in Cole and Morgan (1975) as 'Logic and Conversation'[11] (later published as Chapter 2 of *Studies*):

> 'In the sense in which I am using the word "say", I intend what someone has said to be closely related to the conventional meaning of the words (the sentence) which he has uttered. Suppose someone to have uttered the sentence *He is in the grip of a vice.* Given a knowledge of he English language, but no knowledge of the circumstances of the utterance, one would know something about what the speaker had said, on the assumption that he was speaking standard English, and speaking literally. One would know that he had

[10] Footnote 2 was written with quotes such as this in mind. It's not entirely clear to me whether Clark is characterizing the content of *what is said* or the speech-act of *saying*.

[11] Also published in Davidson and Harman (1972).

> said, about some particular male person or animal *x*, that at the time of the utterance (whatever that was) either (i) *x* was unable to rid himself of a certain bad kind of character-trait or (ii) that some part of *x*'s person was caught in a certain kind of tool or instrument. (Approximate account, of course). 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 (i) and (ii)).'
>
> (II p. 5; *cf.* 1989, p. 25)

Many people building on Grice's saying-implicating distinction have noticed that, actually, this construal of what is said need not exhaust the truth-conditional content of an utterance. As Kent Bach (1994, p. 124) puts it: '...in Gricean terms, the distinction between what is said and what is implicated is not exhaustive.' There has thus been a tremendous amount of work (Wilson and Sperber 1981, Levinson 1989, Sperber and Wilson 1986/1995, Bach 1994, Recanati 1993, Clark 1996, Carston 2002) which uses Grice's distinction as a point of departure from which to chart the territory between the Gricean notion of what is said, construed in the sense of the above quote, and the *broader* notion of truth-conditional content.

The idea that there is some kind of gap between what is said$_{(GRICE)}$ and the intuitive truth-conditional content of an utterance has come to form one part of the *semantic underdeterminacy* (or underdetermination) thesis (see Carston 2002, Chapter One for a comprehensive and engaging presentation of this thesis). In the next sub-section I review some of the ideas behind this thesis, and look at a few of the ways that linguists and philosophers have tried to bridge the gap.

1.2 Underdeterminacy: bridging the gap

Grice's saying-implicating distinction (together with his Co-operative Principle and Maxims) provided the first systematic way of distinguishing what a speaker says from the wider meaning she might intend to convey. Consider cases of metaphor or irony[12] such as those in (3) and (4), or the implicature carried by Jack's reply in (5):

(3) The face of Greece is a palimpsest bearing twelve successive inscriptions.[13]

[*meaning that Greece is rich in history and culture*]

(4) There's nothing remotely interesting in Greece.

[*meaning that Greece is rich in history and culture*]

(5) Lily: I want a holiday somewhere rich in history and culture.

Jack: Have you ever visited Greece?

[*meaning that Greece is rich in history and culture*]

There are, however, a variety of ways in which what is said, construed in the fairly minimal sense of conventional meaning plus reference assignment and disambiguation, falls short not only of the speaker's intended meaning, but also of providing a truth-evaluable proposition at all. Consider (6), (7) and (8):

[12] Actually, it's far from clear that Grice actually has a treatment of either metaphor or irony that goes through. See Wilson and Sperber (2002) for discussion.

[13] From *Travels in Greece*, Nikos Kazantzakis (see bibliography).

(6) Everyone is ready.

(7) Jack drinks too much.

(8) Have you seen Xani's picture?

In (6), there seems little doubt that some process of contextual development distinct from reference assignment and disambiguation must take place before a hearer of this utterance is in a position to establish the proposition expressed. Firstly, in order to specify the domain over which the quantifier is to range there has to be some sort of contextual narrowing of the word 'everyone'. Secondly, it must be clear to the hearer what it is precisely that everyone is 'ready' for—everyone might, for example, be ready to begin the exam, or to listen to the talk, or get on the plane. Likewise, the proposition expressed by an utterance of (7) will include information concerning what it is that Jack drinks too much to do—to operate heavy machinery, for example, or to be entrusted with driving everyone home from the office party. In (8), the precise nature of the relationship between Xani and the picture—e.g. whether it is one that she has painted, one that she likes, or one that she owns—will be part of the proposition expressed by the speaker.

Various proposals have been made as to how this gap might be bridged. One current view is that quantifiers such as 'everyone', expressions such as 'ready', 'too *x*' and the genitive '*s*' construction contribute a hidden, implicit argument—or covert indexical—to the linguistically-encoded logical form of utterances containing them. This provides—in a manner analogous to overt indexicals and pronouns—a linguistic mandate for the contextual development necessary to derive the proposition expressed. A hearer of (6), then, must assign values to

these covert indexicals in order to specify the domain over which the quantifier is to range, find what it is that everyone is ready for, and hence derive the truth-evaluable proposition the speaker is expressing. In example (7), the expression 'too much' sets up a covert slot to be filled (too much [for what?]), as does the genitive construction in (8), which only encodes some, as yet unspecified, relation between Xani and the picture.

Different approaches have different proposals to make as to exactly how a (broadly) Gricean inferential framework might be adapted to assimilate these extra processes. These proposals impact in various ways on how we might conceive of the distinction between semantics and pragmatics, and the relationship between these notions and the notion of what is said. Kent Bach regards them as best analysed as a separate level of meaning *between* what is said and what is implicated. He calls this the level of *impliciture*, and his construal of what is said is necessarily a more minimal, semantic one (see Bach 1994 for further discussion). Francois Recanati, on the other hand, calls them instances of *saturation*, and regards saturation as one of a group of processes that takes place in the derivation of pragmatic aspects of an enriched version of what is said (see Recanati forthcoming).[14] Those working within relevance theory do away with the notion of what is said entirely, and regard the specification of the domain of quantifiers and the enrichment of concepts generally as part of the development of the linguistically-encoded logical form of the utterance into an *explicature*, which they contrast with the level of implicature.

Many different frameworks recognize that the kind of bottom-up, linguistically-mandated processes discussed above are necessary before a hearer

[14] Though crucially, he does not regard these 'primary' pragmatic processes as inferential (see discussion in Chapter Six).

can construct a hypothesis about the proposition expressed by the speaker. More controversial, however, is the existence of cases in which this kind of contextual development takes place in the absence of any covert indexicals or linguistic mandate—i.e. of any slot provided in the linguistically-encoded form. Consider again examples (7) and (8). In (7) most people are happy to interpret the contribution of the verb 'drinks' to the truth conditions of the sentence as one in which it has already been narrowed down to mean 'drinks *alcohol*' (I assumed this implicitly in my two example cases of what Jack might drink too much to do, and doubt that many readers objected). However, it is not hard to think of a context in which it would be narrowed down in a different way. Perhaps, for example, two parents are discussing the reasons why little Rupert keeps wetting his bed. By using the present perfect, a speaker of (8) asks whether someone has seen Xani's picture at some time in the past (*cf.* Jack's question in (5)). But if the speaker of (8) is trying to find the picture Xani has drawn so that she can take it to school, she will be understood to be asking if someone has seen it recently—i.e. since some *contextually-specified* time—most likely in full knowledge that hearer has seen it at some point in the past.

Some, in fact, deny that this kind of *free enrichment* occurs at all (see Stanley 2000, Recanati (forthcoming) and Carston (2000) for discussion). The details are somewhat tangential to concerns here, but I find myself convinced by Carston's arguments that there are good reasons to think that the processes responsible for this kind of contextual development are not linguistically-mandated and, therefore, that the words/expressions in question do not encode hidden indexicals. This is not to say that the development of a linguistically encoded form is unconstrained, simply that the constraints are provided by pragmatic

principles, as opposed to linguistic ones. In the next chapter I explore the extent to which the interpretation of natural behaviours interacts with pragmatic principles and plays a key role in the conceptual narrowing of degree terms such as 'cold', 'happy' and 'angry'.

A final group of cases relevant to the discussion here is one that has been the focus of much attention in relevance theory (Carston 1996, 2002, Wilson and Sperber 1998, 2002). As a kind of parallel process to the narrowing (or enrichment or strengthening) processes that supply the contextual developments discussed above, relevance-theorists also recognize a process they call *loosening* (or broadening or weakening) in which conceptually encoded material might be used less-than-literally. Consider, for example, (9) and (10) below:

(9) The new head-teacher is a fascist.

(10) I hear you've spoken to my secretary.
[*referring to your six-year-old daughter, who has just answered your telephone*]

Although these examples do not strike one as metaphorical in the extended, stylistic sense of (3), there is clearly some contextual loosening process needed to bridge the gap between the literal meaning and the looser interpretation intended by the speaker. To describe a particular head-teacher as a fascist is not necessarily to ascribe to that individual any particular right-wing political views (though of course it might), only to ascribe to them certain fascist-*like* tendencies: perhaps, for example, she is a dogmatic, unreasonable, inflexible disciplinarian. To refer to your seven-year-old daughter as a secretary might, on a

particular occasion, merely be to convey that she has performed a certain secretary-*like* duty (i.e. answered your telephone for you).

Carston, and Wilson and Sperber argue that this kind of loose use of concepts should also be regarded as one aspect of the explicit side of what is communicated by an utterance. It involves the creation of *ad hoc* concepts that are constructed according to considerations of relevance (see Wilson and Sperber 1998, 2002). It is worth noting that aspects of this analysis can indeed be carried over to cases of metaphor such as (3); indeed, it provides an interesting motivation for the very existence of metaphor, which can then be analysed as an extension of the kind of loose use speakers habitually employ.[15]

The debate over free enrichment continues,[16] though it's worth noting that the more weight placed on hidden constituents and semantic incompleteness, the more need there is for powerful, pragmatic principles to guide and constrain the interpretation process (and hence what a speaker can legitimately intend to 'say' or 'have said'). There seems little doubt, however, that there is more to deriving the truth-conditional content of an utterance than merely knowing the conventional meaning of the words uttered, assigning reference and disambiguating ambiguous expressions.

So, to return to questions (A) and (B) from the introduction, we are left with two possibilities. First, we might stop there, and leave it that Grice's notions of saying and what is said are hopelessly flawed, unsuited in crucial ways to the very function that they were designed to perform and hence in serious need of

[15] This approach has implications for any theory of utterance interpretation based on maxims or conventions of truthfulness (such as the one developed by Grice and also Lewis 1967)—see Wilson and Sperber 2002 for discussion.

[16] Despite the 'controversy' I refer to above, there seems little doubt that 'free enrichment' *does* exist. Sentence fragments, for example.

review (if not dismantling entirely). Second, we might return to Grice's writings, and wonder whether he really could have had such plainly unsatisfactory notions in mind. In the next section I explore the second possibility, and suggest he did not.

2. Saying

2.0 Grice's wider programme

In a paper presented in 1991 to the *International Communication Association* in Chicago, Bob Arundale develops a range of carefully researched arguments based on some painstaking comparison of the original *William James Lectures*, the original published versions of some of those lectures (Grice 1968, 1969, 1975), and the revised versions of the original published papers—published as various chapters in *Studies*. He summarizes his main arguments thus:

> 'I argue in what follows that in the well-known 1967 *William James Lectures* Grice had laid out a program with regard to meaning that he had never finished. I argue, as well, that the "new material" in ... *Studies* includes significant clarification of and departure from certain of Grice's earlier positions on issues relating to the concept of meaning. I argue that in view of these "new directions", Grice's 1989 work on meaning can be seen as completing that unfinished program. And I argue, finally, that maintaining these new directions in pursuing Grice's program leads not only to reconceptualising concepts such as timeless meaning, natural versus non-natural meaning, and the said versus the implied, but also to a conception of the Gricean program highly consistent with current research and theory in language pragmatics.'
>
> (1991, p. 2)

My aim here is not to outline the 'new directions' to which Arundale refers,[17] nor is it to assess whether Arundale is right to identify them as such.[18] But I would like to acknowledge here the valuable work he has done, which confirmed some suspicions that originally prompted me to write the paper that forms the basis for this chapter (Wharton 2002).

Instead, I would like to focus on two of the claims made in the above quote.[19] Firstly, I focus on Arundale's claim that in the *William James Lectures* Grice had laid out a wider programme for himself that he never finished. I will then focus on one particular unfinished stage of this wider programme, which I see as highly relevant in any attempt to answer the questions posed in my introduction. The second is Arundale's claim that Grice returns to reconsider this wider programme in *Studies*. This is also pertinent to questions (A) and (B) from my introduction.

---

[17] Although in many ways the third of these 'new directions' is highly relevant to the present chapter. On the strength of his comparisons, Arundale proposes a new definition of *what is said*, and while he openly admits that we can never be sure whether or not Grice would have assented to this definition, he maintains it is consistent with the other new directions he identifies in *Studies* in a way that earlier conceptions of the notion are not. The new definition is: 'Those psychological states that an individual assumes other mature members of his or her language community would generally attribute to a speaker using a given linguistic expression' (1991, p.27). I suspect this condition is necessary, but is hard to see how it would be sufficient; how, for example, can *what is said* in *U*'s utterance of *S* (on this construal of *what is said*) be distinguished from *what is linguistically-encoded* in *S*? Also central to this third new direction is Grice's characterisation of the processes at work behind the conventionalisation of meaning as a 'semi-inferential sequence' (1989, p. 364). I fully agree with Bob Arundale (p.c.) that this direction is worth pursuing, and tentatively address it in Chapter Six.

[18] Arundale tells me (p.c.) that Steve Levinson, for example, once told him that he did not see any 'new directions' in *Studies*.

[19] I should add that I prefer to see what follows (which is, it should be pointed out, a considerably *less* ambitious task than the one Arundale sets himself), in terms of the *re*discovery of an *old* direction.

Whatever programme we do take Grice to have been engaged in at the time of the *William James Lectures*, there was nothing overtly programmatic (or, perhaps better, strategic) about the order in which the lectures were published. For various reasons, the first of the lectures to be published were Lectures VI and VII, which appeared in revised form as Grice (1968), and Lectures V and VI, which were published, again in revised form, as Grice (1969). Stephen Neale offers a plausible explanation:

> 'Although [Grice] produced dozens of first-rate papers, he was always reluctant to go into print—by all accounts heroic efforts were required by editors and friends to extract from him the handful of papers that he deemed worthy of publication.'[20]

As Neale points out, one possible effect of this is that there is a temptation to overlook the points of interaction between Grice's work on saying and his work on meaning.

> 'The *William James Lectures* trickled into print in diverse places between 1968 and 1978 and consequently important connections between the Theory of Conversation and the Theory of Meaning have tended to be missed, ignored or downplayed.'
>
> (1992, p. 556)

Arundale suggests this is a temptation to which some researchers have succumbed:

[20] See *fn.* 2, Chapter One.

> 'Most researchers who have commented on Grice's program on meaning have identified its components in terms similar to Suppes' description (1986, p. 109).'
>
> (1991, p. 6)

The description to which he refers is that offered by Patrick Suppes in his 1986 paper, 'The Primacy of Utterer's Meaning':

> 'As readers of Grice will remember, his programme is to use this basic notion [the fundamental concept of utterer's occasion-meaning] to explicate at the next level of abstraction the concept of utterance-type occasion-meaning. At the next higher level is the analysis of the concept of applied timeless meaning of an utterance type (complete or incomplete) on a particular occasion of utterance. Finally we reach the timeless meaning of an utterance-type...'
>
> (1986, p. 109)

But as Arundale points out, this four-step progression is a somewhat narrower programme than the one Grice had set for himself in the original lectures. Indeed, Grice is explicit on this in the 1968 paper in which aspects of the four-step progression[21] are his main concern:

---

[21] I say '*aspects* of the four-stage progression' because Grice does not actually consider the level of *utterance-type occasion-meaning* in the 1968 paper, so it might be more accurate (when discussing that paper at least) to speak of a *three*-stage progression. It's clear, however, that Grice does not regard assimilating utterance-type occasion-meaning into this three-stage progression as problematic. He remarks in Grice 1969: 'In that paper [1968—TW] I do not distinguish utterance-type occasion-meaning from utterer's occasion-meaning; but once the distinction is made, it should not prove too difficult to difficult to distinguish utterance-type occasion-meaning in terms of utterer's occasion-meaning' (1969, p. 150).

Quote (I)[22]

> 'The account of the (for me) basic notion of meaning is one which I shall not today seek to defend; I should like its approximate correctness to be assumed, so that attention might be focused on its utility, if correct, in the explication of other and (I hope) derivative notions of meaning. This enterprise forms part of a wider program which I shall in a moment delineate, though its later stages lie beyond the limits I have set for this essay.
>
> The wider program just mentioned arises out of a distinction which, for purposes which I need not here specify, I wish to make within the total signification of a remark: a distinction between what the speaker has *said* (in a certain favoured, and maybe in some degree artificial, sense of 'said'), and what he has 'implicated' (e.g., implied, indicated, suggested), taking into account the fact that what he has implicated may be either conventionally implicated...or *non-conventionally* implicated. The programme is directed toward an explication of the favored sense of 'say' and a clarification of its relation to the notion of conventional meaning.'
>
> (1968, p. 225; *cf.* 1989, pp. 117-118)

In the 1968 paper Grice goes on to outline this wider programme, but focuses very much on its early stages, which are, after all, to be the main concern of both that and the 1969 paper. This can be contrasted with an outline of the wider programme that Grice provides in Lecture V of the original lectures:

[22] I'll be referring back to this and the following quotes.

Quote (II)

'Let us take stock. My main efforts so far have been directed as follows:

(1) I have suggested a provisional account of a kind of non-conventional implicature, viz. conversational implicature.

(2) I have attempted to see to what extent the explanation of implicature is useful for deciding about the connection of some of the theses which in Lecture I I listed under the heading of A-philosophical theses.

A lot of unanswered questions remain:

(1) The reliance (without much exposition) on a favoured notion of "saying" which needs to be further elucidated.

(2) The notion of conventional force (conventional meaning) has not received enough attention; (but the notion itself needs to be characterized.)

(3) The notion of conventional implicature needs attention; the relation between what is conventionally implicated and what is said needs characterization.

(4) "Implicature" is a blanket word to avoid having to make choices between words like "imply", "suggest", "indicate", and "mean". These words are worth analyzing.

(5) Also needed: a clarification of the notion of relevance, and a more precise specification of when relevance is expected (filling out the maxim of relevance) and also a further consideration of why there is a general expectation that this maxim (and indeed all maxims) be observed.

I doubt if I shall be able, in the time, to address myself to all of these questions. I shall, in the first instance, try to pursue (1) further: this will carry with it some attention to (2) and (3).'

(V, pp. 12-13 *cf.* 1989, pp 86-87)

He then turns immediately to the first of these unanswered questions,[23] and begins elucidating this favoured notion of saying. Notice that since this lay 'beyond the limits' Grice had set himself for the 1968 and 1969 papers, it is not included in them.

Quote (III)

> 'Saying that p (in the favoured sense of "say"). What follows is a sketch of a direction, not a formulation of a thesis.
>
> (1) I want to say that (A) "U (utterer) said that p" entails (B) U did something x by which U meant that p". But of course many things are examples of the condition specified in (B) which are not cases of saying. e.g. the man in the car; by refraining from turning on his lights he meant that I should go first, he would wait for me, etc.
>
> Let us try substituting:
>
> (2) (C) "U did something x (1) by which U meant that p (2) which is of a type which means 'p'. (i.e. has for some person or other, an established standard or conventional meaning.)" [...]
>
> (3) We want doing x to be a linguistic act; with hideous oversimplification we might try (D)[...]
>
> I will abbreviate this to:
>
> (D') "U did something x (1) by which U meant that p (2) which is an occurrence of a type S which means 'p' in a linguistic system."[...]

[23] 'I doubt if I shall be able, in the time [available?—TW], to address myself to all these questions. I shall, in the first instance, try to pursue (1) further: this will carry with it some attention to (2) and (3).' (V, p.13)

(4) I want to introduce some such idea as that of "central meaning"; I want to be able to explain or talk about what (within what U meant) U centrally meant to give a sense to "In meaning that p, U centrally meant that q".

So "U said that p" may finally come out as meaning:

"U did something x:

(1) by which U centrally meant that p

(2) which is an occurrence of a type S part of the meaning of which is 'p'".

This leaves various questions to be pursued:

(1)[24] How is "U meant that p" to be explicated?

(2) How is "W (word or phrase) means '...'" to be explicated and how is this locution related to "U meant that p"?

(3) How is "S means (would mean) 'p'" (also "S meant 'p' here, on this occasion" and "U meant by S 'p'") to be explicated, how does this relate to the locutions mentioned in (1) and (2)?

(4) How is "U centrally meant that p" to be explicated?'

(V, p. 14 *cf.* 1989, pp. 87-88)

On the strength of the three quotations given above, Arundale summarizes the wider programme Grice sets himself in the original lectures as follows:

'(A) Explicating the concept of what is said, particularly in relation to the concept of conventional meaning, which itself required characterisation. According to Grice's "sketch of a direction", this key step was to be accomplished by means of four, more specific steps:

(A1) Explicating the expression "U meant that p"...or more specifically utterer's occasion meaning.

[24] These are my italics; I'll referring back to these sub-stages of the wider programme (and I'm aware that there are lots of (potentially confusing) numbers in the direct quotation).

(A2) Explicating the expression "W (or S) means '...'", that is, the concept of timeless meaning...and how this is related to utterer's occasion meaning.

(A3) Explicating the expressions "S meant 'p' here" and "U meant by S 'p'", that is, the concepts of applied timeless meaning and of utterance-type occasion-meaning, and how this is related to utterer's occasion meaning.

(A4) Explicating the expression "U centrally meant that p".

(B) Examining the concept of what is "implicated", i.e. the concept of implicature [and] characterizing how conventional implicature is distinct from what is said.

(C) Analyzing terms like "imply", "suggest", "indicate" and "mean" with regard to the concept of implicature.

(D) Clarifying the concept of "relevance" and specifying why there is an expectation that this and other maxims be observed in conversation.'

(Arundale 1991, pp. 7-8)

As I have said, Arundale investigates this wider programme in considerable depth; obviously, however, only certain aspects of it are directly relevant to my concerns here. Stages (C) and (D), for example, do not immediately involve the notions of saying and what is said, and while stage (B) does, I do not intend to elaborate on the (very) brief discussion in Section 2 on the relationship between saying and conventional implicature. The stage of the unfinished programme most relevant to the aims of this chapter (and relevant to the concerns of this thesis) is stage (A)—the explication of the concept of what is said, and it is clear that even this single stage is not completed in the 1968 and 1969 papers, which only deal with sub-stages (A1), (A2) and (A3) (which correlate with questions *(1)*, *(2)* and *(3)* from Quote (III)). It is also clear, and this point seems to me to have been overlooked entirely, that the four-step progression referred to in the

Suppes quote above was intended as a means of *illuminating the notions of saying and what is said*, and is therefore crucial to answering the two questions posed in my introduction.

2.2 The four-step progression

As my source for a rough outline of the four-step progression, or narrower programme discussed extensively in Grice 1968 and 1969, I use a section of Chapter 5 (pp. 88-91 of *Studies*). This section is a considerably revised version of a short section of Lecture VI (p. 11), which appears in the 1969 paper as an introduction to the analysis of utterer's occasion meaning. The basic strategy I will adopt is to briefly consider the different levels of meaning in turn, and pause at each to consider whether they shed any light on the notions of 'say' and 'what is said'; this fits in with what appear to have been Grice's original intentions.

The example sentence—*S*—Grice discusses in the section of Chapter 5 to which I refer is:

'If I shall then be helping the grass to grow, I shall have no time for reading'

The first type of non-natural meaning Grice distinguishes is

- *Timeless meaning(s)*—"*x* (utterance-type)[25] means '....'"

This is the level of conventional meaning(s), or formal content. In the kind of modern parlance adopted in Section 2 this level might be referred to as the level

[25] By which Grice actually had in mind linguistic *or* non-linguistic acts: a sentence or a gesture (hand-signal or headshake).

of linguistically-encoded meaning. The timeless meaning(s) of a *complete* utterance-type such as *S* is (roughly) 'If I shall then be assisting the kind of thing of which lawns are composed to grow, I shall have no time for reading' and 'If I shall then be assisting the marijuana to mature, I shall have no time for reading'. A specification of the timeless meaning(s) of an *in*complete utterance-type—'grass', for example—would be 'the kind of thing of which lawns are composed' and 'marijuana'; in my idiolect at least, another possibility would be 'police informant'.

In the case of pronouns and indexicals, it might help to adapt some terminology from the philosopher David Kaplan (1977/1989), and distinguish between the *character* of a pronoun, and its (truth-conditional/propositional) *content*. So while the content of a given utterance of 'I' will be the person it refers to, the character is the rule (i.e. the timeless meaning) for identifying the referent in any given context. It seems clear that knowledge of the timeless meaning (or character) of *S* would be a necessary condition for knowledge of what *U* says by uttering *S*.

The second type of non-natural meaning Grice distinguishes is:

- *Applied timeless meaning(s)*—"*x* meant *here* '...'"

Fairly transparently, applied timeless meaning is timeless meaning applied to context. Ambiguous expressions have been disambiguated. So in *S*, the applied timeless meaning of 'grass' is 'the kind of thing of which lawns are composed'. In other contexts it might easily be 'marijuana', or 'police informant'. The question of whether reference has been assigned at this level—in other words,

whether this is the level at which Kaplan's content is derived—is harder to answer. Is the referent of an indexical one of its *applied* timeless meanings?[26]

The third type of non-natural meaning is:

- *Utterance-type occasion-meaning*—"*U(tterer)* meant by *x* '...'"

Grice has an extremely subtle distinction in mind here. I quote him at length:

---

[26] This is a complex issue; Schiffer is (not surprisingly) quite clear that pronouns have timeless meaning (or character):

> '... quite roughly speaking, to know the meaning of an indicator word *n* is to know a rule or set of rules such that on the assumption that *Σ(n)* [where '*Σ*' is a complete utterance and *Σ(n)* is read as '*Σ* containing *n*'—TW] was issued with its full conventional force (and assuming one knows the meaning of the other words contained in *Σ* etc.), one will be able to determine... by the application of those rules to the particular circumstances of the issuing of *Σ(n)* what S meant by uttering *Σ(n)*.' (1972, p. 113)

However, he is unwilling to see reference as having been assigned even at the level of utterance-type occasion-meaning, let alone applied timeless meaning.

> '[...] [S]uppose that *S* utters 'He is a conservative' and means thereby that the President of the U.S.A. in 1972 is a conservative. In such an event it would be true to say that *S* said$_i$ that the President of the U.S.A. was a conservative; but it would not be true to say that by 'he' *S* meant "the President of the U.S.A. in 1972".'

If (initially at least) this seems confusing, it is because the 'refers to' sense of 'mean' (as Schiffer points out on p. 5) is one with which he does not concern himself with. Whilst he admits that there will of course be important connections between those senses of 'mean' he chooses to concern himself with and those he will ultimately choose to ignore, it is one that he cannot admit into the framework he builds:

> 'It could perhaps be said that by 'he' *S* meant the President of the U.S.A in1972. This is what might be said in response to the question "Whom did you mean by 'he'?" But here 'mean' has the force of 'refer to', and the expression occurring to the right of the verb 'mean' would not be placed (by me) within double quotation marks.' (1972, pp. 112-113)

In the end we may have to simply abstract away from the problems that indexicals raise for Grice's account generally. Indeed, this is the strategy adopted by Neale 1992.

'It might be true to say that when a particular utterer *U* uttered *S*, *he* meant *by S* (by the words of *S*):

(i) "If I am dead, I shall not know what is going on in the world,"

and possibly, in addition,

(ii) "One advantage of being dead will be that I shall be protected from the horrors of the world".

If it were true to say of *U* that, when uttering *S*, he meant by *S (i)*, it would also be true to say that of *U* that *he* meant *by the words*, "I shall be helping the grass to grow" (which occur within *S*), "I shall then be dead"

On the assumption (which I make) that the phrase "helping the grass to grow," unlike the phrase "pushing up the daisies" is not a recognised idiom,[27] none of the specifications just given of what *U* meant by *S* (or by the words "I shall be helping the grass to grow") would be admissible as specifications of a timeless meaning or of the applied timeless meaning of *S* (or of the words constituting the antecedent in *S*). The words "I shall be helping the grass to grow" neither mean nor mean *here* "I shall then be dead."'

(1989, p. 90; *cf.* 1969, p. 148; VI, p. 11)

Personally, I find assimilating this level of meaning into the overall picture a far from straightforward task. In one way it appears that Grice is describing a level of meaning *between* applied timeless meaning and what is meant (utterer's occasion-meaning). If by the words in *S*, *U* means *(i)* (and *(ii)*), then *by the words* 'I shall be helping the grass to grow' *U* means (utterance-type occasion-meaning) "I shall then be dead". This is clearly a level of meaning distinct from applied

[27] What Grice is getting at here is that since 'pushing up the daisies' *is* a recognized idiom, one of the specifications of the timeless meanings of 'If I shall then be pushing up the daisies...' would be 'If I shall then be dead...' (*cf.* 'If I shall have kicked the bucket...').

timeless meaning: '...the words 'I shall be helping the grass to grow' neither mean nor mean *here* "I shall then be dead"'. In other words, 'I shall then be dead'—a specification of the utterance-type occasion-meaning—is a specification of neither the timeless meaning nor the applied timeless meaning of the words 'I shall be helping the grass to grow'.

To illustrate how this level falls short of the fourth type of non-natural meaning Grice distinguished, it's worth briefly considering that level:

- *Utterer's occasion-meaning*—"*U* meant by uttering *x* that..."[28]

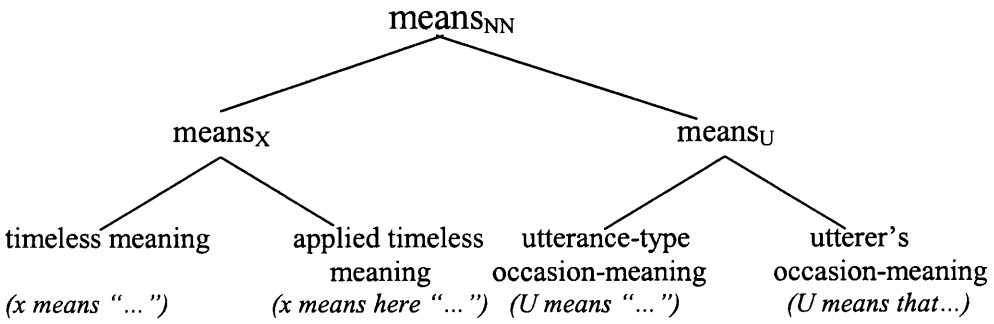
Whilst it may be true that *U* meant—utterance-type occasion-meaning—by the words, 'I shall be helping the grass to grow' (from *S*), 'I shall then be dead', it is certainly *not* true that *U* meant by those words *that* he would be dead (it would follow from this that he was, in fact, dead). Specifications of the utterer's occasion-meaning of *S*, then, would be as follows: '*U* meant by uttering *S* that if he would then be dead, he would not know what is going on in the world'; or '*U* meant by uttering *S* that one advantage of being dead would be that he would be protected from the horrors of the world'.

In another way, however, there is a sense in which it may not be accurate to see utterance-type occasion-meaning as a level *between* applied timeless meaning and utterer's occasion-meaning. In the introductory chapter to his 1972 book *Meaning*, Stephen Schiffer sets out the various notions of meaning he (following Grice) is working with. He begins by introducing two basic senses of $mean_{NN}$: '$means_S$'—as in 'He meant something by that remark', and '$means_X$'—

[28] The absence of the quotation marks is crucial.

as in 'That word means something'. I have illustrated the further distinctions he then draws in *fig. 2* below:[29]

fig. 2



Seen in terms of the above diagram, it is less obvious that the level of utterance-type occasion-meaning is indeed a level *between* applied timeless meaning and utterer's occasion meaning. Rather, it is a necessary level of abstraction in order that 'means$_X$' and 'means$_U$' can be separated; recall that for Grice, ultimately, the former was to be analysed in terms of the latter, so each node on the bottom of the diagram is to be analysed in terms of the notion(s) of meaning corresponding to the nodes to its right.

As I said, the intuitions behind utterance-type occasion-meaning are by no means straightforward. According to Schiffer, utterances of the form "*U* means '...'" are used 'to report the sense or meaning of *x* [*U*] intended to have (or be

[29] A few notes on the diagram: firstly, Schiffer does not use exactly the same terminology as Grice, but the distinctions he draws appear to me to be roughly parallel (though it should be noted that Schiffer's analysis of 'means$_X$' is considerably more complex than my over-simplified diagram suggests); secondly, Schiffer uses *double* quotation marks to enclose *what is meant.* Grice (though he is inconsistent—compare the way in which he sets out the four levels in 1989, p. 90-91, with his discussion of utterance-type occasion-meaning on page 90) tends to use single quotation marks (I don't think much rides on this); thirdly, in *Meaning*, Schiffer uses '*S*' to denote the speaker. To avoid confusion, and since I have been using '*S*' to signify the *sentence* and '*U*' to signify the speaker (or utterer), I use *my* abbreviation here.

operative) on the occasion of his producing (or doing) *x*' (1972, p. 3). The problem, as far as I can see, is not that that it is hard to see what an occasion-specific meaning might be, but rather how, on a certain occasion, it might distinguished from the fourth level of meaning. In other words, where utterance-type occasion-meaning would *end.*

At this point, one question that seems to me to be at least in the spirit of reconciling the narrower programme outlined in the 1968 and 1969 papers with stage (A) of the wider programme set out above (and the problem outlined in the previous paragraph) is the question of whether or not Grice would have regarded knowledge of the utterance-type occasion meaning of an utterance of *S* as a necessary part of knowing what *U* had *said* in an utterance of *S*. By uttering *S*, does *U say (i)* or *(ii)*?

Regarding *(ii)*, it seems highly unlikely: the words in *S* take the form of a conditional, and the words in *(ii)* do not. Grice, in fact, appears to confirm this in the section of the original lectures of which the above quote is a considerably revised version. There (but not, for some reason, in the revised section in Chapter 5 of *Studies*), he writes:

> 'When he said *S*, *X* [*U*—TW] meant that:
>
> (a) Since he would be dead, he would know nothing of what is happening in the world; and possibly <u>also</u> (in context):
>
> (b) One advantage of being (as he would be) dead would be that he would be protected from the horrors of the world; (*though maybe one would prefer that he implied that* (b)).'[30]
>
> (VI, p. 11—my italics, TW)

[30] Grice continues this passage: 'Perhaps corrections, etc. are included under "means", while "taking it further" would be included under "implies".'

But does *U*, by uttering *S*, *say (i)*? Since none of the discussion of the four types of non-natural meaning Grice distinguishes in Chapter 5 of *Studies* actually features the word 'say', all this is largely a matter of interpretation. But notice that if it were the case that Grice regarded knowledge of utterance-type occasion-meaning—i.e. *(i)*—as a necessary part of knowing what *U* said in an utterance of *S*, then the interpretation of 'say' (and the derivative notion what is said) would be crucially different to the minimal sense(s) outlined in Section 2. My own interpretation is that Grice *would indeed* have regarded *U* as saying *(i)* (and regarded knowledge of *(i)* as a necessary condition for knowing what is said by (or in) *U*'s utterance of *S*), but I will leave discussion of my reasons until the next sub-section, where I hope to equip myself with the right kind of (hopefully, watertight) terminology to keep myself afloat in some pretty perilous philosophical waters.

In the next sub-section I look at the second of Arundale's claims, that Grice returns in *Studies* to reconsider the wider programme. Since, as I said, I am abstracting away from stages (B), (C) and (D), this will involve considering sub-stage (A4), which Grice did not address in either the 1968 or the 1969 paper.

2.3 Riddle of the strands: dictiveness and dictive meaning

As I said, in reworking Lectures VI and VII for the 1968 paper, and Lectures V and VI for the 1969 paper, Grice chose to omit the outline he provides in quotes *(II)* and *(III)*. Interestingly, though, he chose to re-insert it in the revised version of the 1969 paper published as Chapter 5 of *Studies*, and included both quotes in an introduction—newly entitled 'Saying and Meaning'—to that chapter. (It should also be noted that the material Grice selected for Chapters 5 and 6 reflects

the chronology of the original lectures, rather than the date of publication of the revised versions of those lectures.) Of course, we can only speculate on why Grice omitted these sections from the original published versions and then chose to reintroduce them into *Studies*; Arundale speculates as follows:

> '...between the 1968 and 1969 articles, and the new work on meaning included in *Studies* in 1989, Grice did not return to the issues of the wider program outlined in William James Lecture V. Given this observation, one is led again to wonder why he decided to reinsert the same initial outline as the introduction to Chapter 5 of *Studies in the Way of Words*. I maintain that in reinserting the outline, Grice was refocusing attention on the issue of explicating what is said, and thereby providing a frame for his moves in the Retrospective Epilogue...toward completing the unfinished wider program.'
>
> (1991, p. 9)

This seems to me to be a possibility worthy of investigation. It is certainly true that while Grice (1968) and (1969) address sub-stages (A1), (A2) and (A3) of stage (A), neither of the papers gets as far as sub-stage (A4). Of the various questions Grice addresses in beginning his attempt to elucidate saying in quote *(III)*, questions which he raises as a result of the 'unanswered questions' remarked on in quote *(II)*—questions which *themselves* should be seen in the light of the overall aim of stage (A)—only *(1)*, *(2)* and *(3)* are dealt with in any detail in the 1968 and 1969 papers. Other than the few lines devoted to the latter stages of the wider programme in the outline provided in the 1968 paper, question *(4)*—'How is "U centrally meant that p" to be explicated?'—is not addressed at all.

In Strand Five of his Retrospective Epilogue (1989, pp. 339-385), Grice returns to consider it. Some of his comments here might be construed as being at odds with the minimalist view of what is said attributed to him on the basis of the quote from Lecture II above (p. 80); they might even be interpreted as evidence of inconsistency in his work. What I would like to suggest, however, is that we go at least some of the way to resolving this apparent inconsistency if we interpret these final thoughts bearing in mind that Grice was returning to consider aspects of the wider programme outlined in the original lectures.

Grice's concern in Strand Five is to discuss two candidates for:

> '...any kind, type, mode or region of signification which has special claims to centrality, and so might offer itself as a core around which more peripheral cases of signification might cluster, perhaps in a dependent posture.'
>
> (1989, p. 359)

The two candidates are *dictiveness* and *formality*. Grice is quite explicit (see 1989, pp. 360-361) that the former is to be taken as equivalent to what is said, and the latter as equivalent to conventional meaning. He then goes on to outline various ways in which the two notions can be shown to be mutually independent. So as well as the existence of utterances that have both dictive and formal content, and those that have neither dictive nor formal content, we find utterances in which only one of the two candidates is present. Utterances containing the kind of non-truth conditional connectives considered briefly in Section 2—'but', 'moreover', 'on the other hand'—would contain as a component formal, non-dictive content (+ formal – dictive), in virtue of the fact that (according to Grice's analysis) these expressions indicate the performance of non-central

speech acts, and their contribution to the overall speaker meaning is not part of what is said.

As a further consideration, and as further evidence of the dissociation between dictiveness and formality, Grice then discusses example utterances where the dictive content might be regarded as 'nonformal' (– formal + dictive):

> 'Suppose someone, in a suitable context, says "Heigh Ho". It is possible that he might thereby mean something like "Well that's the way the world goes".[31] Or again, if someone were to say "He's just an evangelist", he might mean, perhaps, "He is a sanctimonious, hypocritical, racist, reactionary, money-grubber". If in each case his meaning were as suggested, it might well be claimed that what he meant was in fact what his words said; in which case his words would be dictive but their dictive content would be nonformal and not part of the conventional meaning of the words used. We should thus find dictiveness without formality.'
>
> (1989, p. 361)

Given the interpretation of 'what is said' discussed in Section 2 (and based on the quote on p. 52 above from Lecture II), this is puzzling. What exactly does Grice mean by dictiveness, and how can it be reconciled with the construal of what is said outlined in Section 2?

I will assume that Grice was not a person to introduce new terminology without good reason. We can, for example, speculate why he chose to use the

[31] For those who think that I am crediting Grice with notions of 'saying' and 'what is said' that extend too far beyond what Grice (allegedly) envisaged in Lecture II (1975, 1989: Chapter 2) (i.e. by suggesting earlier that Grice would indeed have regarded *(i)* as being said in *U*'s utterance of *S*), I think the 'Heigh Ho' example is highly suggestive that I am not. I should add that, actually, I would not go along with Grice's analysis of 'Heigh Ho'.

term 'formal' as opposed to 'conventional' in these, his final comments. From early on in his work (Grice 1957, reprinted as Chapter 14 of *Studies*) Grice is seeking alternatives to the word 'conventional':[32]

> 'This question about the distinction between natural and non-natural meaning is, I think, what people are getting at when they display an interest in a distinction between 'natural' and 'conventional' signs. But I think my formulation is better. For some things which can $mean_{NN}$ are not signs (e.g. words are not), and some are not conventional in any ordinary sense (e.g. certain gestures).'
>
> (1989, p. 215)

Regarding *dictive* content or *dictiveness*, however, there are no clues either in the original publications of the lectures (1968, 1969, 1975 or 1978), nor in the revised papers in *Studies* (Chapters 6, 5, 2 and 3 respectively). Again, though, the original typescripts of the lectures are illuminating.

In the final lecture (Lecture VII) of the *William James Lectures*, Grice provides a detailed outline of his progress in addressing the wider programme outlined in Lecture V. This section of Lecture VII appeared in revised (and more finely detailed) form as the outline for the wider programme presented in Grice 1968. Points (4) and (5) are highly significant:

> 'Recapitulation. In Lecture V I set before myself the task of doing something to elucidate my notion of "saying" and my use of "conventional meaning" ("conventional force"). This enterprise resolved itself into...
>
> (1) [...]

[32] See *fn.* 9.

(2) [...]

(3) [...]

(4) An attempt to identify the core of conventional meaning, or more exactly an attempt to define "U dictively meant that *p" where the specification of what U dictively meant need/would not exhaust the account of what U (conventionally) meant. To say "U dictively meant that *p" is a first approximation to a defniniens for "U said that *p".

(5) An attempt to show how elements in what U meant, which do not belong to what U dictively meant, get in (how they are relative to U's dictive meaning). A specification of at least one class of "conventional implicatures" (hopefully).'

(VII, p. 1 *cf.* 1968, pp. 225-229)

The crucial point to note here is that Grice identifies 'dictively meaning that p' with 'saying that p' (or at least a first approximation of it). Not only, then, does Grice return in Strand Five to reconsider sub-stage (A4), but he returns to terminology he had used in setting out the wider programme in the original lectures to do so.[33]

[33] Arundale points to what he considers to be some inconsistencies in Grice's thinking here. If, he claims, Grice's aim 'in examining 'centrality' was to aid in distinguishing what is said from what is implied..., then to identify one of the candidates for centrality as the presence or absence of 'dictive content' is patently tautological because the candidate concept is simply another label for the distinction between the said and the implied!' (1991, p. 26). Also, Arundale claims, 'Grice first introduced the phrase "dictively meant" in place of the phrase "centrally meant"...at the start of Lecture VII [...]. Evidently, Grice himself saw dictive meaning and what is said as synonymous, so that examining dictiveness as one of the candidates for centrality that might elucidate what is said would not appear to be a productive move' (1991, p. 26). Regarding the second claim, it's not immediately clear to me that Grice did regard the notion of dictive meaning as *synonymous* with that of central meaning—it appears to me that Grice was seeking to refine the notion, rather than merely give it another name; I would admit that the relationship is rather unclear, but I think Grice had good reason (though he did not give it at the time) to use the term

Arundale points out that in Lecture VII Grice used the term 'dictive meaning' 'without explanation'. What I will offer is (I hope) an explanation; in this regard I will suggest that there are similarities between Grice's notion and a notion introduced by another post-war Oxford philosopher—R. M. Hare—and, furthermore, that that these similarities extend to more than mere etymology.[34]

In his (1949/1971) paper *Imperative Sentences*, Hare attacks the view that only a certain class of sentences—summarized in the following quote—are "the proper subject-matter of logic":

> 'The sort of sentences which are to be admitted into the logical fold are variously referred to as "scientific", "cognitive", "informative", "fact-stating", "true-or-false", "theoretical", "referential", "symbolic" etc.; and the sort of sentences which are to be excluded are called "emotive", "evocative", "non-fact-stating" etc. The latter are held not to state genuine propositions, and therefore, since propositions are the bricks out of which a logical system is built, to be altogether beyond the pale of such a system. They are sometimes even said to be "literally senseless.'"
>
> (1971, p. 1)

Hare does this by arguing that logic can be used to shed light on imperative sentences too. Essentially, his argument goes, imperative sentences can also be

---

'dictive meaning' rather than 'central meaning' in Lecture VII, and, equally, return to the term 'dictiveness' in Strand Five; this I explore below.

Also, we shouldn't forget that Grice himself admits that, given the breadth of the programme on which he is embarking in the lectures, his exposition includes moments of 'hideous oversimplification' and he can only offer 'a sketch of a direction, not a formulation of a thesis'. The fact that, nearly 40 years on, Grice's work is still so rich and rewarding suggests that Grice's 'sketch of a direction' is worth more than most researchers' highly detailed maps.

[34] Grandy and Warner (1986, p. 7) also refer to the ties between Grice's and Hare's work.

shown to have properties that enter into logical relations. These logical relations hold because indicative and imperative sentences share similar properties.

I adapt my example sentences in (2bc) in order to illustrate the distinctions Hare introduces to demonstrate these shared properties. Consider the indicative and imperative sentences in (14ab):

(14a) Zoë goes to nursery.

(14b) Zoë, go to nursery!

For Hare, both these sentences have the same *descriptive* content (or, as he later termed it, *phrastic*), represented below as ($14ab_{desc.}$). The part of a sentence that performs the descriptive function is known as the *descriptor*, which, as can be seen, is identical to the speech-act notion of descriptive content or propositional content.

($14ab_{desc.}$) Zoë goes to school at time *t*.

Following Hare's original mode of presentation, an alternative way in which the difference in meaning between (14a) and (14b) might be represented is by adding the words 'yes' and 'please' to the descriptive content, as in ($14a_{dict.}$) and ($14b_{dict.}$):

($14a_{dict.}$) Zoë goes to school at time *t*, *yes*.

($14b_{dict.}$) Zoë goes to school at time *t*, *please*.

'Yes' and 'please' in these sentences have no other function than to indicate the mood of the sentence and the illocutionary force with which the sentence is uttered. Hare calls this function the *dictive* function, and he calls words that perform this function *dictors*. For Hare, it is the dictor that does the saying; the descriptor simply specifies what it is that is being said. In most languages, Hare points out, dictors are implicit: 'Even in English, however, we can say of a sentence what mood it is in; there must, therefore, be something about it which tells us this. This, then, is the dictor and the rest is the descriptor' (1971, p. 9).

In an appendix to the chapter in Hare (1971) in which the 1949 paper is published, Hare discusses the relationship between his notion of dictor (indicative mood, assertive force) and the 'assertion-sign' as used by Frege and Russell. Whilst on the face of it the two notions appear to be the same, Hare points out that there are important differences (1971, pp. 22-24). Frege and Russell, for example, assume tacitly throughout their work that all sentences are indicative, and their 'assertion-sign' only serves to indicate subscription to a particular proposition, i.e. to *assert* that it is true.[35] This, as Hare points out, is one, but not the only, function of the dictor (Hare's counterpart to the assertion-sign), which serves to indicate *mood* as well as force.

In his later work (1970/1971) Hare refines his dictor/descriptor distinction and takes account of this by separating the dictor into two other notions: the *neustic* and the *tropic*.[36] The function of the tropic is to indicate mood, as distinct from *subscription* to the phrastic (or descriptive) content, which is indicated by

[35] Recanati (1987) discusses this in detail, and how it might shed light on Dummett's (1981) interpretation of Frege.

[36] He also introduces a separate notion—*clistic*—to represent 'completeness', but this is not directly relevant to the discussion in hand.

the neustic. Thus, as an illustration, when a declarative sentence is embedded in the antecedent of a conditional, or appears in direct quotation marks, whilst retaining its indicative tropic, it does not have the requisite neustic, since it appears unasserted. As far as Hare is concerned, however: 'Neustics are normally understood with uttered sentences unless something special is done to indicate that they are not being subscribed to.'(1970/1971, p. 91)

Grice was clearly aware of Hare's distinctions; in fact, in the latter part of Strand Five he makes explicit mention of 'philosophers...who, in one way or another, have drawn a distinction between "phrastics" and "neustics"' (1989, p. 367): presumably, as the originator of the distinction, Hare is one of these. It seems plausible to suggest, then, that when Grice equates what is said with what is *dictively* meant, he is surely doing so with the notion of Hare's *dictor* (i.e. neustic and tropic) in mind.

It is clear that for Grice 'saying' and speaker commitment (or, more accurately, those psychological attitudes involved in a speaker's commitment) were inextricably linked in a way Grice saw as absolutely central to providing any satisfactory characterization of 'saying'. He remarks on this in Lecture III:

> 'It is not a natural use of language to describe one who has said that *p* as having, for example, "implied", "indicated", or "suggested" that he believes that *p*; the natural thing to say is that he has <u>expressed</u> (or at least purported to express) the belief that *p*. He has of course committed himself, in a certain way, to its being the case that he believes that *p*; and while this commitment is not a case of <u>saying</u> that he <u>believes</u> that *p*, it is bound up, in a special way, with saying <u>that *p*</u>.'
>
> (III, p. 2; *cf.* 1989, p. 42)

A detailed discussion of Hare's distinctions, and the relationship between them and Austin's locutionary/illocutionary distinction, Peter Strawson and John Searle's reinterpretation of it, and Hare and Jonathan Cohen's arguments contra-Austin, is to be found in the section of Recanati (1987) referred to in *fn.* 8. One passage from Recanati's discussion brings Grice's views on the relation between saying and speaker commitment (and the notion of dictive meaning) neatly into focus. Recanati is presenting Hare's phrastic-tropic-neustic distinction and how it might relate to saying and asserting:

> 'In example (6)...a single proposition is expressed twice, at first asserted then mentioned, without this being indicated by quotation marks:
>
> (6) — You are an imbecile.
> — Oh, I am an imbecile! Thank you very much!
>
> ... In (6), in particular, the [*second—TW*] speaker "says that" he is an imbecile, but does not subscribe to what he says. His utterance is a case of "saying that"...by virtue of its declarative tropic, but it is not a bona fide assertion—it lacks what Hare calls the neustic.'
>
> (1987, p. 263)

But for Grice, the second speaker does not 'say' *anything* (or at least not this): saying entails meaning, and the speaker clearly does not mean here that she is an imbecile. The fact that the second speaker's utterance lacks the neustic *precludes* it from being a case of 'saying that' for Grice. 'Saying' *must* involve a 'neustic', otherwise it would not be meant (and hence could not be said).

It is this, I believe, that lies behind Grice's use of the terms 'dictive meaning' and 'dictively meant' in Lecture VII, and 'dictiveness' in Strand Five. At the level of applied timeless meaning, ambiguous expressions have been disambiguated (and, possibly, reference has been assigned), but *this is not enough to exhaust what is said.* Before *S* can be regarded as having been said by *U*, *S* must have been *stated* by *U*. In her utterance of *S* 'If I shall then be helping the grass to grow, I shall have no time for reading' *U* says 'If I am dead, I shall not know what is going on in the world' because that is what *U* states or dictively means; parallel to this, in her utterance of "He's just an evangelist" *U* dictively means 'He is a sanctimonious, hypocritical, racist, reactionary, money-grubber'. For Grice, what is said is what is stated: it is not just the proposition expressed by the speaker, it is *the proposition expressed by the speaker and presented as true.*

Sperber and Wilson (2002) discuss Grice's notion of saying in some depth. They point out that if Grice really had in mind a notion of 'say' that did *not* involve speaker commitment, it is hard to see how his analysis of tropes, for example, would have worked at all:[37]

> 'Understood in this way, the maxim of truthfulness means "Do not express propositions you believe to be false." The function of this maxim, and more generally of the Quality maxims, would be to account for the fact—to the extent that it is a fact—that a speaker actually commits herself to the truth of what she says. Tropes would then be explained by the claim that flouting the maxim triggers the recovery of an implicature in the standard Gricean way. However, there is a problem. In general, the recovery of implicatures is meant

[37] To be fair, they also point out the shortcomings of a view in which 'saying' equates with 'expressing as true'; that is, that it appears to render the first Quality maxim—the Maxim of Truthfulness—redundant.

> to restore the assumption that the maxims have been observed, or that their violation was justified in the circumstances (as when a speaker is justified by her ignorance in providing less information than required) (Grice 1989, p. 370). In the case of tropes, the maxim of truthfulness is irretrievably violated, and the implicature provides no circumstantial justification whatsoever.'
>
> (Wilson and Sperber 2002, p. 589)

Returning to question (A), then, what Grice meant by 'saying' was 'stating'. Naturally, it follows from this that what Grice meant by what is said (question (B)) was what is stated.

If knowledge of the utterance-type occasion-meaning of *S* does indeed contribute to knowing what *U* had said in an utterance of *S*, it might go some way to solving what is in some way one of the biggest mysteries in Grice's notion of dictiveness without formality: how might dictive content be 'nonformal', and hence not part of the conventional meaning of the words uttered? Just as by uttering *S*, *U* might mean something like *(i)* above, *U* might also mean by uttering 'He's just an evangelist', 'He's just a sanctimonious, hypocritical, racist, reactionary, money-grubber'. As with the notion of utterance-type occasion-meaning, the idea in Strand Five is that this is what *U does* with the words: the content of the central act—the *saying*—performed by the speaker of those words. Notice, crucially, that this meaning does not necessarily form part of the conventional meaning (timeless and applied timeless meaning) of the words, though it is *closely enough related to* the conventional meaning[38] to be part of what is said. In both the 'grass to grow' and the

[38] Arundale is not impressed with Grice's example, and clearly does not attach as much importance to it as I do. '[It] purported to show dictive content without formality, but did so only

'evangelist' examples, what is said clearly goes beyond conventional meaning, and it might thus be inferred that what we have is 'dictiveness' *without* 'formality'.

A related point to the claim that what is said must include Hare's neustic, and one which is highly relevant to any attempt to answer Question (B), is that in many of Grice's formalisms, it is *not just* the propositional (descriptive, phrastic) content which is embedded under 'saying that'. Grice writes in Lecture VII, for example, that he is attempting to find a definiens for "U said that *p"[39] or "U dictively meant that *p" (VII, p. 1). According to these formulations it is quite clear that what is said *includes* what Grice refers to as a 'dummy' mood indicator: '*': what is said, then, includes Hare's 'tropic', as well as his 'phrastic'.

This suggests another way on which Grice's conception of what is said might diverge, not only from the minimalist view sketched in Section 2, but also from the kind of speech-act approach championed by John Searle also discussed in that section; in fact, it might thus line up better with John Austin's original version of the locutionary/illocutionary distinction than Searle's version of it as a

---

on the grounds that a special language convention shared by a small group is not the convention shared by the larger population' (1991, p. 26). I, however, prefer to see these occasion-specific senses (either the example of utterance-type occasion-meaning from 1989, p. 90, or the example of [– formality + dictiveness] from Strand Five) as *nonce*-senses, constructed 'on-the-fly', as it were. I don't believe for a second that Grice thought this could only be achieved in virtue of a convention shared by a small group of language-users; that would surely be [+ formality] anyway. Grice saw communication as a rational activity to be characterised (ultimately) as a *meeting-of-minds*, not mere adherence to convention (a notion which, recall, he was unhappy with anyway).

[39] In his John Locke Lectures delivered at Oxford in 1979 Grice uses the term *radical* to refer to the 'phrastic'—i.e. '*p*'. According to Dummett (1981) the term 'radical' is originally due to Wittgenstein; it refers to 'sentences considered as having sense but not force' (1981, p. 328).

distinction between descriptive content and illocutionary force (see Recanati 1987 for discussion).

It could be argued that Grice's is a slightly 'loose' use of the expression 'mood indicator'. If this is so, however, it is by no means one that was untypical of him. Richard Warner remarks (Grice 2001, p. *xiv*) that during the *Immanuel Kant Lectures* Grice delivered at Stanford in 1977, Julius Moravscik objected to Grice's use of the word 'mood', on the grounds that he was using it in a sense other than the standard linguistic sense. As can be see from the discussion so far, it's certainly clear that Grice did not have the standard *syntactic* sense of mood in mind—in which it simply means some kind of verbal inflection—at all, and that he intended it in a broader, *semantic* sense, as including those semantic properties that distinguish standard declarative utterances, for example, from interrogative or imperative ones.

In his later work (see the *John Locke Lectures* delivered at Oxford in 1979—and published as *Aspects of Reason* in 2001) Grice makes clear that he is actually working with an entirely different conception of mood. So much so, that (presumably in the light of Moravscik's objection) he ceased calling *his* notion 'mood' at all):

> 'I am ... talking about "modes" rather than "moods" to make it clear that I am not trying to characterize what linguists would be likely to call "moods" (though I would expect there to be important links between their "moods" and my "modes"). I would justify (or explain)[40] my use of the term "mode" by reference to my views about meaning. According to these views, *what a speaker means* is to be

[40] Grice's idiosyncratic use of parentheses is (so I have found) one that is (highly) contagious.

> explained in terms of the effect which he intends to produce in an actual or possible hearer; and *what a sentence in a language means* is to be explained in terms of directives with respect to the employment of that sentence, in a primitive (basic) way, with a view to inducing in a hearer a certain kind of effect; [...] The intended effect on a hearer is (in my view) one or other of a set of psychological attitudes with respect to some "propositional content" (to borrow momentarily a phrase I do not normally use), and my mode-markers each correspond with one element in this set of attitudes (or set of "modes of thinking")."
>
> (2001, pp. 68-69)[41]

This aspect of Grice's approach raises interesting questions concerning Grice's view of the relationship between language and thought. For whereas Hare saw 'tropics' and 'phrastics' as properties of sentences, Grice appears to have been moving in another direction entirely. As is clear from the above quote, mode (mood), was a matter of *thought*.[42] Regarding the 'phrastic', or—as Grice called it—the 'radical', Grice is quite explicit: 'I regard it as an undecided question whether there are any sentences in a natural language which contain a part which is a distinct surface counterpart of a radical' (2001, p. 50).

One possible objection to a proposal that what is said contains a mood indicator is that it is difficult to get an intuitive grasp on how what is said, so construed, would look. In fact, this may be one of the reasons why most theorists have taken mood to be somehow external to propositional content. Notice,

---

[41] This quote is taken from a section in which Grice is attempting to account for the difference between practical and alethic modals in terms of mode.

[42] The project Grice that is engaged in during this later work—broadly speaking, an enquiry into *reason* and *rationality*—also ties up in interesting ways with Grice's work on meaning and conversation; indeed, according to Grandy and Warner 1986 it represents the *third stage* of Grice's work on meaning.

however, it does not follow from the fact that what is said includes mood that Grice required each individual clause of the propositions (or representations) that form the vehicles of thought to have their own mood 'built in', as it were. Recall that for Grice, even basic semantic notions such as sentence meaning were ultimately to be characterised in terms of propositional-attitude psychology. Under this approach, *semantic* mood is as derivative a notion as syntactic mood, to be analysed ultimately in terms of the manner in which a particular thought is entertained, rather than the surface appearance of the language with which a speaker may choose to express it.

3. Implications

Reassessing Grice's notions of *saying* and *what is said* in this way has a number of implications, a few of which I have attempted to address at the appropriate stage of the presentation above. However, there are two others that I would like to discuss briefly here.

The first concerns cases of metaphor. Grice regarded these as cases of categorial falsehood. So when *U* utters 'Beckham's right boot is once more the bow that delivered the fatal arrow', the first Maxim of Quality is flouted and a related comparison is implicated. But if I say of an atheist acquaintance of mine 'He's just an evangelist', and by uttering those words the sense or meaning of *S* I intend to have conveyed on the occasion of producing *S* (i.e. the utterance-type occasion-meaning) is 'He is just a sanctimonious, hypocritical, racist...etc.' (and this is what I say), then am I not in some sense at least also uttering a falsehood (since, presumably, he is not an evangelist)? In cases such as these, how can a

line be drawn between occasion-specific senses and what is implicated? Just how closely related to the conventional meaning of a word (or sentence) does utterance-type occasion-meaning have to be. Here I return to the issue I mentioned early in Section 3 of where utterance-type occasion-meaning might end.

Carston (2002) draws interesting parallels between Grice's examples in Strand Five and the relevance theory treatment of 'loose use' in terms of *ad hoc* concepts mentioned briefly in Section 2:

> '[T]he intended interpretation [of 'He's just an evangelist'—TW] has the characteristics of the sort of loose use of lexical concepts that I'm interested in here. It seems to involve a dropping of the encoded (conventional) PREACHER concept and the picking out of a set of relevant encyclopaedic properties: HYPOCRITICAL, REACTIONARY, etc.'
>
> (Carston 2002, p. 330)

So stepping back, and surveying the scene from a more current, cognitivist perspective, we might say that the encoded linguistic content of *U*'s utterance of 'He's just an evangelist' would be (16):

(16) He_x is just an evangelist.

But Grice had intuitions that what *U* 'says' in his utterance of this sentence is what she asserts. So what is stated—including applied timeless meaning and utterance-type occasion-meaning—is (17):

(17) He$_x$ is an evangelist* (i.e. 'sanctimonious, hypocritical, racist...etc.')

Where, in a cognitive sense, EVANGELIST* is an *ad hoc* concept constructed on the fly, as it were, or in Grice's sense—[– formal + dictive]—word meaning. Since as I mentioned in Section 2, relevance theorists extend this treatment to metaphor, it is interesting to speculate that Grice might have been led eventually to propose that metaphor worked in a similar way (and hence differently from irony).[43]

The second major implication should be clear: and it concerns the semantic underdeterminacy thesis. If 'saying' is co-extensive with 'stating', then what a speaker says is co-extensive with what a speaker states.[44] But what is *stated* in examples (6)-(10) (repeated below as (18)-(22)) has to be a proposition in which the gap between the minimal notion of what is said introduced in Section 2 and the intuitive truth-conditions has already been bridged.[45] If not, *U* does not *mean* what *U* utters (and hence does not *say* anything at all):

---

[43] Schiffer (1972, p. 111) considers this issue of the relationship between utterance-type occasion-meaning and metaphor: 'Notice, too, that in order for *S* [*U*—TW] to mean "..." by *x*, *S* must think that *x* means "...".[...] There are apparent counter-examples to this feature of our definition, but I believe that they are not actual counter-examples. For example, *S* may utter 'Richard is a lion' and mean thereby that Richard is brave. In such a case one might be inclined to say that *S* meant "brave" by 'lion'. However, I should want to say that one who said this would either be speaking incorrectly or speaking elliptically.' My intuition is that *U* does indeed mean "brave" by 'lion', and that she is speaking neither incorrectly nor elliptically, simply—as most speakers do—loosely.

[44] I think I'm in good company here: '(I should like to add that, as far as my intuitions go, in uttering σS stated (asserted) that *p* just in case in uttering σS said that *p*.)' (Schiffer 1972, p. 114).

[45] It also includes a mood indicator.

(18) Everyone is ready.

(19) Jack drinks too much.

(20) Have you seen Jack's picture?

(21) The head-teacher is a fascist.

(22) I see you've spoken to my secretary.

In (18), for example, what is stated must include the domain over which the quantifier is to range and, furthermore, for what it is precisely that 'everyone' is ready. Parallel arguments can be constructed for (19)-(22). In fact, the last two examples (particularly (21)) are highly suggestive of the examples Grice himself provides for utterance-type occasion-meaning and dictiveness without formality. Since saying is stating, what is said cannot be, as Herb Clark remarks, simply 'the literal meaning of the sentence uttered with its ambiguities resolved and its referents specified', it is a *fully truth-evaluable proposition*. For Grice, the 'gap' simply did not exist, and the minimal notion of what is said so often attributed to him *cannot* be what he meant.

Exactly how Grice expected hearers to derive those aspects of what said that go beyond conventional meaning is unclear. A natural solution, of course, would be his Cooperative Principle and Maxims themselves. Neale (1992, p. 530) uncovers the following quote (from Grice 1957), in which Grice hints that the Maxims (or, more accurately, factors that the Maxims were eventually designed to account for) may indeed play a role in inferring the speaker's intentions behind an utterance containing an ambiguous expression:

> '[I]n cases where there is doubt, say, about which of two or more things an utterer intends to convey, we tend to refer to the context

> (linguistic or otherwise) of the utterance and ask which of the alternatives would be relevant to other things he is saying or doing, or which intention in a particular situation would fit in with some purpose he obviously has (e.g. a man who calls for a 'pump' at a fire would not want a bicycle pump). Nonlinguistic parallels are obvious: context is a criterion in settling the question why a man who has just put a cigarette in his mouth has put his hand in his pocket; relevance to an obvious end is a criterion in settling why a man is running away from a bull.'
>
> (1989, p. 222)

But in the end (see 1989, p. 375) Grice seems strangely reluctant to make what seems like a fairly natural step.

It should be noted, however, that this issue is a problem even for those who maintain the kind of minimal notion of what is said discussed in Section 2. 'Context' (undefined) is often regarded as sufficient to assign reference and disambiguate ambiguous expressions (and hence derive the minimal notion of what is said), but it is clear in the above quote that Grice regards disambiguation certainly as a process of working out which reading of an ambiguous word a speaker *intends* to convey; it is far from clear how this might take place without some (surely) pragmatic machinery by which these intentions might be recovered.

Having said that, among the tantalizing clues that Grice left there are some which suggest that he does seem to have had in mind at least some role for the maxims in determining what is said. Ironically enough, given the concerns of this thesis, the point he raises involves natural behaviours of the kind that are the central theme here. Apart from the discussion of frowning in his 1957 paper, Grice devotes little attention to such phenomena; however, in Lecture III

(published as Chapter 3 of Studies) Grice discusses 'natural' features of speech—stress and intonation. Wielding 'modified' Occam's razor as ever, Grice declares he is reluctant to attribute conventional (or timeless) meaning where not absolutely necessary, and hence is reluctant to attribute it to, for example, stress. He goes on:

> '...we might first introduce a slight extension to the maxim enjoining relevance, making it apply not only to what is said but also to features of the means used for saying what is said. This extension will perhaps entitle us to expect that an aspect of an utterance which it is within the power of the speaker to eliminate or vary, even if it is introduced unreflectively, will have a purpose connected with what is currently being communicated, unless, of course, its presence can be explained in some other way.'
>
> (III, p. 15; 1989, p. 51)

Notwithstanding this, the purpose of his chapter has *not* been to make the wider claim that Grice saw a role for his Maxims in deriving what is said, simply to reassess what Grice meant by say (and derivatively, what is said). Regarding this, I hope I have at least revealed the prospect that views often ascribed to Grice, views that are sometimes presented in the literature as cornerstones of the traditional Gricean approach, may well turn out to be ones that he never held. And yes, I realize there's a great deal of *interpretation* here, and that a great deal of the above depends on a very personal view of what one takes Grice to have meant by what he said and wrote, but since that is where we came in, and since it is *Grice* we are talking about here, I hope I am justified in doing so.

In the next chapter I present the framework adopted in this thesis—relevance theory. I look more closely at the relation between showing and saying and

suggest that rather than the dichotomy Grice envisaged, there is, in fact, a continuum of cases between the two (though I will stress that the continuum actually runs between cases of showing and cases of *meaning*$_{NN}$). This has considerable bearing on not only the analysis of natural behaviours I will offer, but also, ultimately, on what should be seen as the domain of pragmatic principles or maxims; for it suggests that they are best seen as applying to the domain of overt intentional communication as a whole, rather to the domain of meaning$_{NN}$: precisely the approach relevance theory takes.

I also explore some of the consequences of adopting a fully inferential model of communication. I illustrate the central role played in relevance theory pragmatics by the inferential adjustment of linguistically encoded content, and show how natural behaviours contribute at an explicit, as well as an implicit level. In other words, as Grice recognizes in the quote above, these behaviours not only play a role in an audience understanding what a speaker meant, but also in understanding what a speaker has said. I also look at different types of encoding that are made possible within a relevance-theoretic framework.

Chapter Three

1. Relevance theory and the showing-meaning$_{NN}$ continuum

'There is a point where too much information and too much information processing can hurt. Cognition is the art of focusing on the relevant and deliberately ignoring the rest.'

(Gigerenzer and Todd 1999, p. 21)

In his *William James Lectures*, Grice proposed that human verbal communication is a co-operative activity driven by the mutual expectation that participants will obey a *Cooperative Principle* and *Conversational Maxims* of *Quantity*, *Quality*, *Relation* and *Manner*.[1] In outlining his theory of conversation, one of Grice's main aims was to cast light on some of what he regarded as the 'illegitimate applications' (1989, p. 3) of certain philosophical 'manoeuvres' of members of the ordinary language philosophy movement. This was a movement that had influenced him greatly at Oxford in the nineteen-forties and fifties, a movement that (arguably) he belonged to, and which, in opposing the central tenets of the 'idealised' language philosophy of Frege, Russell and Carnap, was instrumental in the birth and development of modern pragmatics.[2]

---

[1] Grice notes that 'the specific expectations or presumptions connected with at least some of the... maxims have their analogues in the sphere of transactions that are not talk exchanges' (1989, p. 28). He goes on to illustrate other 'transactions' (two people mending a car, or baking a cake) in which participants might reasonably expect some degree of co-operation from others along the lines of those suggested by the cooperative principle and maxims.

[2] The term 'pragmatics' is not indexed in *Studies*, nor is it mentioned much by Grice in either the original lectures or the subsequent published versions of those lectures (or other published work). However, Grice does use the word in Strands Seven and Eight of the Retrospective Epilogue of *Studies* when discussing his reactions to (what he calls) Strawson's 'Neo-Traditionalism'—that is, Strawson's reaction to Russell's 'Modernism': 'A few years after the appearance of *Introduction to Logical Theory* I was devoting much attention to what might loosely be called the distinction between logical and pragmatic inferences. [...] ... I canvassed the

Relevance theory (Sperber and Wilson 1986/1995, Wilson and Sperber 2002) builds solidly on Gricean foundations. It combines aspects of Gricean intention-based pragmatics with modern research in psychology and cognitive science to provide a cognitive-inferential framework. Natural language is seen as governed by a code, itself governed by an autonomous mental grammar. Utterance interpretation, on the other hand, is a two-stage process. The linguistically encoded logical form, which is the output of the mental grammar, is simply a starting point for rich inferential processes guided by the presumption that speakers will conform to certain expectations of relevance. In contrast with conscious, reflective reasoning, these inferential processes are unconscious and fast, under-pinned by 'fast and frugal heuristics' of the kind currently gaining much currency in cognitive science (see Gigerenzer, Todd and the ABC Research Group 1999).

Relevance theory is built around two principles. The *Cognitive Principle of Relevance* makes a fundamental assumption about human cognition: the human cognitive system is geared to look out for *relevant* information, which will interact with existing mentally-represented information and bring about positive cognitive effects based on a combination of new and old information. Relevance itself is a property of inputs to cognitive processes, and is defined in terms of cognitive effects gained and processing effort expended: other things being equal, the more cognitive effects gained, and the less processing effort expended in

---

idea that the alleged divergences between Modernists' Logic and vulgar connectives might be represented as a matter not of logical but of pragmatic import' (1989, p. 375). Given that Russellian 'Modernism' was one of the targets of much of the work undertaken by the ordinary language philosophers, there's an interesting irony in the fact that Grice's 'pragmatics' was conceived largely in an effort to *defend* it.

gaining those effects, the greater the relevance of the input to the individual who processes it.

The human disposition to search for relevance is seen as an evolved consequence of the tendency toward greater efficiency in cognition (Sperber and Wilson 2002).[3] In Dan Sperber's words:

> 'Cognitive efficiency involves making the right choices in selecting which available new attention to attend to and which available past information to process it with. The right choices in this respect consist in bringing together input and memory information, the joint processing of which will provide as much cognitive effect as possible for as little effort as possible.'
>
> (1996, p. 114)

It is, furthermore, a disposition that is routinely exploited in human communication. Since speakers know that listeners will pay attention only to stimuli that are relevant enough, in order to attract and hold an audience's attention, they should make their communicative stimuli appear at least relevant enough to be worth processing. More precisely, the *Communicative Principle of Relevance* claims that by *overtly* displaying an intention to inform—producing an utterance or other *ostensive stimulus*—a communicator creates a presumption that the stimulus is at least relevant enough to be worth processing, and moreover, the most relevant one compatible with her own abilities and preferences. Recall from the last chapter that the motivation for making the kind of 'less direct' inferences I discussed there is the very fact that a communicator has created in her audience

[3] The evolutionary function (in the sense of the fitness-enhancing effect) of human cognition is to provide individuals with information about themselves and their environment (and thus guide behaviour). See Sperber 2001 for discussion.

an expectation that there is something worth their while to infer. Relevance theory is an attempt to flesh out the notion of what makes communicated information worthwhile.

As has already been pointed out, in contrast to Grice's (1957) aim of characterising meaning$_{NN}$, relevance theory aims at providing a characterisation of human overt intentional communication generally. Utterances, after all, are not the only kind of ostensive stimuli, and a communicator might provide evidence of her intention to inform by means of a look, a gesture, or—as we saw in Chapter One—even a natural sign. Ostensive stimuli are, more often than not, a mixture of what Grice would have called natural *and* non-natural meaning, and this is one of the reasons that relevance theory does not attempt to draw the line that Grice wanted to between "'deliberately and openly letting someone know" and "telling"' (1989, p. 218).

Recall the characterization in Chapter One: in any act carried out with the intention of revealing an informative intention, there are two layers of information to be retrieved. The first, basic layer is the information being pointed out, and the second is the information that the first layer is being pointed out intentionally. What makes an individual ostensive act a case of either 'showing' or 'meaning$_{NN}$' is the precise nature of the evidence provided for the first layer. In cases of showing, the evidence provided is relatively direct—Schiffer's bandaged leg, for example. In cases of meaning$_{NN}$, the evidence provided is relatively indirect—a linguistic utterance, for example. Sperber and Wilson (1986/1995, p. 53) discuss the relationship between the two notions:

> 'Is there a dividing line between instances of ostension which one would be more inclined to describe as "showing something", and clear cases of communication where the communicator unquestionably "means something"?... What we have tried to show... is that there are not two distinct and well-defined classes, but a continuum of cases of ostension ranging from "showing", where strong direct evidence for the basic layer of information is provided, to "saying that",[4] where all the evidence is indirect...'

As I have suggested, this has clear implications for what should be seen as the domain of pragmatic principles or maxims, for it suggests that they are best seen as applying to the domain of overt intentional communication as a whole, rather to the domain of $meaning_{NN}$. Relevance theory, then, recognises both showing and $meaning_{NN}$ as instances of overt intentional or—as they term it—*ostensive-inferential* communication. Most cases of showing—cases in which the evidence provided is fairly direct—*still* require an extra layer of inference before the

[4] Earlier versions of the continuum I am about to propose were dubbed the 'showing/*telling*' continuum. I now prefer 'showing/$meaning_{NN}$' continuum—though the reader will notice I do occasionally lapse into the 'showing/*saying*' continuum, another earlier version.

The reason for the change in terminology is that the 'showing/telling' continuum and the 'showing/saying' continuum both fail to acknowledge that there exist cases of $meaning_{NN}$ which are cases of neither saying nor telling. The reason for my occasional inconsistency is that as my thinking for this thesis has developed, it has become clear to me that whereas I originally thought there was one, there are actually *two* continua—between *x* and *y*, and between *w* and *z*. (These two continua are conflated in, for example, Wharton 2000b, 2003a.) Indeed, I have subsequently come to realise that this was one of the reasons I had problems conceptualising the continuum in my earlier work (a fact which was never clearer to me than when I tried to explain it to others). The issue is further complicated by the fact that *one* of these two continua, the kind of continuum proposed by people like Erving Goffman (1981) and Adam Kendon (1988), might *indeed* be better called the 'showing–*saying*' (or 'saying *that*') continuum—in the sense that it is more a continuum between non-linguistic coding and linguistic coding, than a continuum between showing and $meaning_{NN}$. In Chapter Four I explore the similarities between the two types of continua, while in Chapter Six I stress the differences between them (whilst also noticing that they interact in interesting ways).

communicator's full informative intention is recognised[5] (recall that in example (6d) from Chapter One (p. 39) Mary's mother still has to make the less direct inference that Mary wants help), and the extent to which an audience is required to make this extra inference is a *question of degree.*

Consider Scenario One below in the light of Grice's photograph example (see pp. 32-33 above):

> *Scenario One:*
>
> I am a private detective, hired by Mr. *X* to follow Mrs. *X* (he suspects that she is having a relationship with Mr. *Y*). I have taken a photograph of Mrs. *X* and Mr. *Y* together; the quality is poor (I used a telephoto lens, and there is a little camera-shake), and a blurred image of the couple can only just be seen in the distance (though on close inspection it is unmistakably them).

There is a subtle, but to my mind clear, difference between my photograph example and Grice's original one. As far as I can see, it is this. If I leave *my* photograph in Mr. *X*'s room by accident, it is no longer absolutely clear that Mr. *X*'s coming across the photo will induce the same effect on him as my showing it to him. It is only by close inspection that he could even see this was a photo of Mrs. *X* and Mr. *Y*. Crucially, then, it may only be in virtue of my showing it to him that Mr. *X* would take the time and effort to look at the photograph closely enough to make out exactly who the photo shows. In other words, whether or not the photograph has the effect I desire may well depend on Mr. *X*'s successful

[5] Notice that an audience is only required to *recognise* a speaker's informative intention; he might not believe the speaker, in which case the informative intention will be recognised but not fulfilled. By contrast, the communicative intention may be *fulfilled* without being recognised, in that it can be evidenced without the audience consciously attending to it.

recognition of my intention to produce some effect on him *by means of the recognition of that intention.*[6]

Of course, it might be responded that the degree to which Mr. *X* is required to attribute intentions to me in this scenario is minimal. Nonetheless, the requirement is there, and it seems clear that the recognition of my intention to inform Mr. *X* does indeed play some role, however minimally, in accounting for the effect of my photograph on him.

Consider also Scenario Two:

> *Scenario Two:*
> I am a private detective, and also a keen amateur photographer. I have taken another (better) photo of Mrs. *X* and Mr. *Y*, and I have developed it myself. As a photographer, I am proud of the colour, the contrast and the general quality of the print. I show a friend of mine the photograph of Mr. *Y* displaying undue familiarity to Mrs. *X*.

How does my friend respond? In the scenario provided so far, the tendency is to suggest that my friend would probably remark something like 'My goodness, Mr. *Y* is certainly having an affair with Mrs. *X*', or even 'I hope you're going to show this to Mr. *X*'. Suppose, however, that I tell you the friend referred to above us is a professional photographer. Suddenly, a variety of other responses may be appropriate: 'The colour is great' or ' I love the quality of the light' or 'Aren't

[6] Sam Guttenplan (p.c.) has remarked to me that, in his experience, undergraduates are not usually convinced by Grice's photograph example. He suggests that this is because when Grice wrote the original article, photographs were more clearly strict causal products of reality. Nowadays, modern digital photography and computer image manipulation complicate the issue somewhat, and we more readily call the authenticity of the (purported) causal link into question.

those new polarisers terrific?' And what does my colleague's response depend on? *What he takes my intention to be in showing him the photograph.*

Actually, I think the point can be made even more clearly. You are walking down the street and a complete stranger comes up to you and thrusts a photograph in front of your face. Having recovered from the initial shock, most people would probably react in the same way: with utter confusion. Of course, like me, you might say 'Lovely!' or 'How interesting!' but that would only be because you thought that by responding in such a way you might get rid of this mad photograph-shower. Actually, I bet most people wouldn't know what to think, the problem being that although it would be perfectly clear that you were being shown a photograph, it would be far from clear exactly what it was you were being shown a photograph *of* (or what you were being shown the photograph *for*).

Although it requires a slight stretching of the imagination, even in Grice's original photograph example (the comparison between (1) and (2) on pages 32-33 above) is there not a sense in which Mr. *X* must attribute to the photograph-shower the intentions *behind* his showing it? It will, after all, make a difference to the effect of the picture on Mr. *X* whether or not he takes the shower to be intending to inform him about Mrs. *X* and Mr. *Y*, and not to be just showing him the quality of the colour, or the light, or the new polariser he has invested in. As Deirdre Wilson pointed out to me in conversation, even if two individuals *A* and *B* are in the same room as two other people—for the sake of convenience Mrs. *X* and Mr. *Y*—engaged in (as she put it) 'unfaithful activities', there will still be at least some degree of intention-attribution involved if *A* attempts to point out something about them to *B*: other things being equal, you might just as easily be

pointing out something Mr. *Y* is (or isn't) wearing, as drawing attention to the (potential) inappropriateness of their behaviour. Not only, then, can what is $meant_{NN}$ only be regarded as a *sub-set* of what is intentionally communicated, but rather than the dichotomy Grice envisaged in his 1957 paper, there is a continuum of cases between showing and $meaning_{NN}$.

It might be objected that we are running the risk of allowing into the domain of pragmatic principles or mechanisms all manner of cases in which what is communicated is either so weak or so vague that it cannot be adequately characterised. The response, I suggest, is that overt communication often *is* weak and vague, and that a theory of human communication should at least try and accommodate these vaguer aspects. We saw in Chapter One that there are examples which are clearly cases of intentional communication that do not qualify as cases of $meaning_{NN}$ according to Grice's definitions; we now see that—even if we wanted make a distinction—there is no convenient cut-off point between the two where such a distinction might be drawn. If we attempt to limit our attention to cases that are uncontroversially cases of $meaning_{NN}$, then we are forced to ignore a whole range of communicative exchanges that deserve explanation.

To help account for the vaguer aspects of communication, including the communication of impressions, emotions, attitudes, feelings and sensations, Sperber and Wilson propose that the informative intention might be better characterised as an intention to modify not the hearer's thoughts directly, but his *cognitive environment*: this includes not only all the facts or assumptions that he is aware of, but also all the facts or assumptions he is capable of becoming aware of, in his physical environment—in relevance-theoretic terms, the set of facts that

are manifest to him (i.e. that he is capable of perceiving or inferring). This notion of *manifestness* plays a central role in the relevance-theoretic characterisation of an *informative* intention, which is defined not in Gricean terms, as an intention 'to produce a particular response *r*', but rather as an intention 'to make manifest or more manifest to the audience a set of assumptions **I**' (p. 58).[7]

An assumption may be manifest to different degrees. The more salient a manifest assumption, and hence the more likely to be mentally represented, the more strongly it is manifest. Vague communication typically involves a marginal increase in the manifestness of a very wide range of weakly manifest assumptions, resulting in an increased similarity between the cognitive environments of communicator and audience.

Consider the following example. Jack and Lily have returned by ferry to the Greek island on which they first met. They disembark. Having scanned the quayside, he smiles at her; then he looks back ostensively to the quayside again, urging her to look too. She gazes along the quayside. What is Jack drawing her attention to? Is it the *taverna* at the water's edge, the octopus drying in the breeze, the ragged cats sniffing the nets, the bougainvillea in the *kastro*, the brilliant light? Is it one, many or *all* of these things?

---

[7] The notion of manifestness is also central to the relevance-theoretic notion of a *communicative* intention, defined as an intention 'to to make it mutually manifest to audience and communicator that the communicator has [an] informative intention' (p. 61). The psychologically plausible notion of mutual *manifestness* is the relevance theory solution to the problems with Grice's original three-clause definition of meaning$_{NN}$ raised by (among others) Strawson 1964 and Schiffer 1972. (See Avramides 1989, Chapter 2 for discussion of Strawson's and Schiffer's counterexamples and redefinitions; see Sperber and Wilson 1986/1995 Chapter 1 for further discussion.)

But Lily does not turn to Jack and say 'What do you mean?'. She acknowledges him and smiles back, because she understands. The sights, sounds and smells perceivable in her physical environment interact with her inferential abilities and her memories to alter her cognitive environment, making it possible for her to have further thoughts, memories and feelings similar to his own. This is all that Jack intended: to convey an impression. He did not *mean* anything; his intention cannot be pinned down to one specific proposition or small *set of propositions*; it was simply to make more manifest to Lily whatever assumptions became manifest to him as he scanned the quayside.

On other occasions, when the intention might be to communicate something equally intangible, and equally hard to spell out in words—emotions or feelings—it might also be preferable to use a behaviour that falls somewhere between showing and meaning or saying. Given the vagueness of some of the ostensive stimuli that constitute cases of 'showing', it seems clear that this intention is not always reducible to an intention to communicate simply a single proposition and propositional attitude (or even a small set). In the next chapter I will argue that interjections such as *aha*, *wow* and *ouch* are often used to communicate in similarly vague ways, marginally increasing the manifestness of a very wide range of assumptions.

The continuum, then, has a variety of applications. At various points along it, we can see the varying extents to which hearers are required to consider intentions of speakers in order to get from the evidence they provide to the first, basic layer of information they are communicating. It therefore provides a 'snapshot' of the types of evidence used in intentional communicative acts and the role inference plays in them. At one extreme of the continuum lie clear cases

of spontaneous, natural display; at the other extreme lie clear cases of linguistic coding, where all the evidence provided for the first, basic layer is indirect. In between lie a range of cases in which more or less direct 'natural' evidence and more or less *in*direct coded evidence mix to various degrees: for example, in pointing and stylised expressions of emotion. Equally importantly, the continuum provides a theoretical tool which allows us to conceptualise more clearly the observation made above that ostensive stimuli are often highly complex *composites* of different, inter-related behaviours which fall at various points between 'showing' and 'meaning$_{NN}$'.[8]

In the next section I discuss a further consequence of a shift to a relevance theory framework. In many instances, it seems, what we tend to regard as the vaguer, natural aspects of the complex ostensive stimuli mentioned above play a central role in verbal communication: not only do they aid the audience's determination of what the speaker has meant$_{NN}$, but they also contribute to determining what the speaker has *said*.

[8] Although the focus in this thesis is very much on what I am calling natural 'behaviours', there are clearly all manner of other ways in which natural signs can be (and are) used in verbal communication. If, while sitting behind my desk, I am approached by an undergraduate, who asks me if I am willing to play in the annual Linguistics *vs.* Phonetics rugby match, I might reply 'I've broken my leg' whilst at the same time drawing attention to the fact that the leg in question is in plaster. If, while attending a particularly infuriating meeting, I decide to register my disapproval by closing by brief-case and walking out, I might close my brief-case and walk out, and at the same time utter 'I'm closing my brief-case and walking out'.

2. Relevance theory and lexical pragmatics

> 'If my husband were to tell me he was disappointed he had missed me at lunch, I would wonder if he meant to say he was sad—which is simply regretfully sorry; unhappy—which is somewhere between mad and sad; mad—which makes you want to argue with someone over what they had done; angry—which makes you want to ignore the person you are feeling this way towards; furious—which makes you want to spit; or none of the above. In order for me to really understand what people are saying I need much more than a few words mechanically placed together... Words by themselves are too vague.'
>
> (Liane Holliday Willey 1993, p. 63)

One of many parallels between relevance theory and the pragmatic framework developed by Grice is that relevance theory distinguishes between the *explicit* and the *implicit* content of an utterance. This distinction bears some similarity to Grice's famous distinction between *saying* and *implicating*, the distinction which—together with his Co-operative Principle and Maxims—provided the first systematic way of distinguishing what a speaker says from the wider meaning she might intend to convey.

However, notwithstanding the discussion in the previous chapter, the two pairs of notions are not entirely parallel (see Carston 2002 for in depth discussion). In relevance theory, *explicatures* are recovered via a mixture of linguistic decoding *and* inference and are also a matter of degree: the greater the degree of linguistic encoding, the *more explicit* the communicated content of the utterance. Compare three ways that Lily might reply to Jack's question "Do you like the gift?":

(1a) (smiling happily) I think it's wonderful. I like it very much and I feel absolutely delighted.

(1b) (smiling happily) I do.

(1c) (smiling happily) I've always wanted an electric toothbrush.

What Lily encodes in (1a) is an incomplete logical form that Jack will decode and develop inferentially into a fully propositional form which will constitute the basic explicature of the utterance. In (1b) what she encodes is very fragmentary indeed, and Jack is left to do considerably more inferential work to identify the explicit content (or explicature) of her utterance. In (1c) what she encodes is a conceptual representation that Jack must not only develop into an explicature, but complement with an implicature in order to derive an answer to his question.

A central claim of relevance-theoretic pragmatics is that explicatures and implicatures are developed in parallel, with the explicit content being adjusted or 'fine-tuned' in various ways in order to yield the implicatures required to satisfy the audience's expectations of relevance. In particular, the encoded content of lexical items occurring in an utterance may have to be narrowed or loosened (assigned a narrower or broader denotation) in order to yield the expected level of implicatures. Given the arguments of relevance theory, and in particular the characterisation of 'expectations of relevance' mentioned above, there is a straightforward procedure which may be used to determine the appropriate degree of narrowing or loosening, and more generally, to identify the speaker's 'meaning'. I will illustrate this with an example. Consider Jack's utterance of (2) below, as he talks to a friend on the telephone while watching Lily opening her gift:

(2) She's opening the parcel.

Given that the stimulus Jack has provided is presumed by the hearer to be at least relevant enough to be worth processing and, moreover, the most relevant one compatible with Jack's abilities and preferences, the hearer is justified in following a path of least mental effort in looking for the intended effects, which should make the utterance relevant enough. This may involve assigning reference and disambiguating, narrowing or loosening lexical meaning, and supplying particular contextual assumptions in order to derive the expected level of effects. Once his expectations of relevance are satisfied, it is reasonable for him to conclude that the 'meaning' he has inferred was the one the communicator intended.[9] The following comprehension procedure—taken from Wilson and Sperber (2002, p.13)—is therefore rationally motivated:

Relevance theoretic comprehension procedure

(a) Follow a path of least effort in computing cognitive effects: Test interpretive hypotheses (disambiguations, reference resolutions, implicatures, etc.) in order of accessibility.

(b) Stop when your expectations of relevance are satisfied.

In order for the hearer to derive the explicature of (2), he must assign reference to the pronoun 'she' and consider the most accessible candidate first. Following a path of least effort he will form the hypothesis that Jack is referring to Lily. He is also likely to narrow the concept encoded by the verb 'opening'—which denotes a quite general action that may be performed in various ways—and take Jack to mean that Lily is opening the parcel in a certain way, e.g. untying the ribbon or undoing the sellotape, perhaps in a way characteristic of

[9] The processes of hypothesis formation and evaluation that an audience undertakes are non-demonstrative in character. So he may be *wrong*, but he can do no better.

her. This can be contrasted with the sense of 'opening' conveyed in (3) and (4) below, which would have quite different implications:

(3) The dog is opening the parcel.

(4) The blue tit is opening the milk-bottle.

In (3) the sense of 'open' conveyed would more likely involve the tearing of paper (and gnashing of teeth); in (4) the sense conveyed is totally different again, involving as it does pecking through the foil top on a bottle of milk. Given the relevance-theoretic comprehension procedure, the hearer of (3) can take it that if Jack had intended to convey that Lily was 'opening' the parcel in a non-standard manner (one less easy to imagine and hence more costly in terms of effort), he would not have chosen to convey this with the utterance he used.

What this suggests is that there is an interaction between decoding and inference not only the level of what is explicitly communicated by a whole sentence, but at *word* level too: a particular word may be used to express not exactly the concept it encodes, but another related concept—constructed by drawing on encyclopaedic information—which is more relevant in a given context (see Sperber and Wilson 1998, Wilson 1998 for discussion of verbs such as 'open'; see also Carston 2002, Wilson and Sperber 2002). In such cases, the hearer constructs an *ad hoc* concept guided by considerations of relevance along the lines sketched in the previous chapter.

Consider the word 'bear', the concept for which might be *narrowed* in (5) to denote a sub-set of bears (e.g. polar bears), or *loosened* in (6) to include objects which are not strictly bears at all (e.g. large hairy dogs):

(5) The bear walked out across the frozen sea.

(6) I loved Emma, my Old English Sheepdog: she was a bear.

This kind of adjustment of conceptual content is a feature of relevance-theoretic pragmatics, and is seen as one of the principal processes in explicit communication. A point that has been little remarked on is that the interpretation of natural communicative phenomena ('paralinguistic' features) feeds directly into the lexical adjustment process. Consider examples (7), (8) and (9) below:

(7) Jack: Shall we sit out here?

Lily (shivering ostensively): I'm *cold.*

(8) Lily (furiously): That makes me *angry*!

(9) Lily (smiling broadly): I feel *happy.*

In (7), Lily and Jack meet outside a café. Lily's ostensive shiver accompanying her utterance of 'I'm cold' is picked out by the relevance-theoretic comprehension procedure and used in interpreting the degree term 'cold'. It will be treated as commensurate with the degree of coldness she feels, and, in effect, will *calibrate* the degree of coldness Jack understands her to feel and to be expressing as part of her meaning. The fact that Lily has shivered ostensively—*shown*, as well as *told* him she is cold—motivates Jack's search for the 'extra' meaning Lily intends to convey. Clearly in this case, implicatures may depend on it; thus, Jack might be entitled to infer that Lily is definitely cold enough to want to go inside. In a parallel example, Lily's ostensive shiver accompanying her utterance of 'It's lovely out here on the terrace, isn't it?' might provide Jack with

a clue that she is being ironic, that actually she hates it on the terrace and that she would prefer to go inside. In both cases, shown natural behaviours feed into the interpretive process, guiding the hearer to a certain range or type of conclusions.

Notice also, however, that the natural behaviours not only help Jack establish the implicit content of Lily's utterance, but also *the proposition he takes Lily to be expressing* (or the explicature of her utterance). The truth conditions of her utterance of 'I'm cold'—and the truth-conditions of (7) and (8), which also contain degree terms—will vary according to the type or degree of 'coldness' (or 'anger' or 'happiness') she intends to communicate, and hence reflects in her natural behaviour.

What Lily linguistically encodes by using the word 'angry', for example, is some quite general concept, which encompasses a considerable range of degrees and types of anger accessible via encyclopaedic (contextual) assumptions. The linguistically encoded content is calibrated by Lily's furious tone of voice and enriched by Jack to a concept—ANGRY*—that he sees as commensurate with the degree and type of anger Lily intends to convey. What she encodes by the use of the word 'happy' is also a quite general concept;[10] again the occasion-specific sense is calibrated by reference to Lily's natural behaviours—in this example by features of her smile. The process is relevance-driven, and the *ad hoc* concepts are constructed along the same lines as in examples (2), (4) and (5).

Relevance theory also distinguishes between the proposition expressed by an utterance (or the basic-level explicature) and a range of *higher-level explicatures*. These are constructed by embedding the basic truth-conditional content under a speech-act or propositional-attitude description, which may be explicitly

[10] Though see discussion in *fn.* 11.

indicated or pragmatically inferred. Consider utterances (10a) and (11a) below, which would lead a hearer to construct the higher-level explicatures in (10b) and (11b):

(10a) Regrettably, your application has been unsuccessful.

(10b) The speaker regards it as regrettable that my application has been unsuccessful.

(11a) Frankly, you haven't got the job.

(11b) The speaker is telling me frankly that I haven't got the job.

Notice also, however, that this kind of attitudinal information can also be conveyed by entirely natural behaviours. So a speaker of (10a) might convey her attitude by speaking in a regretful tone of voice, and a speaker of (11a) might convey the fact that she is speaking frankly simply by adopting a frank manner.

In everyday communication we simply take for granted how a speaker naturally displays a certain degree of emotional intensity or attitude, and how (equally naturally) an audience has the ability to discriminate subtle variations in tone of voice and facial expression. In the case of human paralinguistic behaviours, certainly, a great many appear to work along *analogue* lines, and in this way can be directly contrasted with the *digital* code of language (I return to this analogy in Chapter Five). We read natural signs much as the engineer studies the needle on an analogue pressure gauge, in which the needle's movement is analogous to the rising and falling of the pressure, and continuous pressure fluctuation is reflected in the continuous movement of the needle. In Lily's utterance, her frown and angry tone of voice are in a similarly proportional or

analogous relationship to the amount of affect he intends to convey. Depending on the gravity of her frown and the tone of voice she uses, Jack might decide she is mildly annoyed, quite angry or absolutely furious. The extent to which Jack can interpret these degrees of her annoyance or anger or happiness depends not on his knowledge of any digital code, but on his ability to discriminate among tiny variations in her facial expression and tone of voice, much as the engineer reads the quivering needle.

As I say, in our everyday conversational exchanges, we take the natural side of verbal communication for granted, but if we are trying to construct a pragmatic theory, we should not (and in cases of Asperger's syndrome, as the quote at the head of this section shows, we often *can* not). The role of the more 'natural' aspects of complex ostensive stimuli in establishing speaker's 'meaning'—including basic and higher-level explicatures—should be neither overlooked nor downplayed.

3. *Translational* and *non-translational* activation of concepts

> 'The type of co-ordination aimed at in most verbal exchanges is best compared to the co-ordination between people taking a stroll together rather than to that between people marching in step...'
>
> (Sperber and Wilson 1998, p. 199)

Broadly speaking, what a communicator *C* wants to communicate to her audience *A* is a (more or less complex) thought. I take it, following Fodor (1983) and others, that thought takes place in some sort of modality-neutral representational medium, or 'language of thought'. This neutral medium takes the form of structured sets of concepts: *conceptual* representations. These representations have logical properties, and are capable of being true or false. As a result, a

conceptual representation can contradict or imply other conceptual representations, and act as input to logical inference rules. Inference, after all, just is the manipulation of conceptual representations under truth-conditional (or evidence-based) constraints.

If humans were truly telepathic and genuinely able to read each others' minds, *C* would simply 'put' her thought into *A*'s mind (or *A* would somehow retrieve it). Humans, however, are not telepathic, and in order to communicate her thought, *C* must produce some publicly observable behaviour which will get *A* to entertain the same (or, more accurately, a similar) thought to her own. The hedge here is important, for there are two ways of cashing out the phrase 'communicate a thought'. According to a code model of communication, *C*'s thoughts are translated into a signal by use of a code, and translated back into the original message by *A*. As we shall see in subsequent chapters, this kind of account may work well in the case of bee-dancing, but it does not adequately characterise human linguistic communication.

One effect of a shift to an inferential model is to allow for the possibility that the thought *A* entertains as a result of communicating with *C* may only *resemble* *C*'s original thought, rather than duplicate it. It is this observation that is behind the 'looseness' to which Sperber and Wilson refer in the above epigraph. Paradoxically, the code model of communication results in a duplication of thoughts similar to the kind one would expect if humans were indeed telepathic. Sperber (2001) calls it 'cognition by proxy', and argues that the fact that human communication is inferential has interesting implications for how we account for its evolution. I return to issues such as this in my final chapter, and suggest it has implications for what we take to be the function of communication.

That being said, there is of course a coded element to human linguistic communication: some words *do* translate into concepts, the constituents of thoughts. If I utter the words 'open' or 'bear' to a competent English speaker (i.e. someone who knows the code), the appropriate concepts will be activated in his mind; the same goes for a competent Portuguese speaker on hearing 'abrir' or 'urso'. We might call this the *translational* activation of concepts, and the kind of coding that gives rise to it *translational coding*. Within an inferential model, as we saw in the last section, there is room for these concepts to be narrowed or loosened along lines already explored; nonetheless, these processes can only take place once the concept has been (translationally) activated.[11]

However, as we saw in the last chapter, and as speech-act theorists notably recognised, linguistic meaning need not be of just one type. In cognitive terms, this raises the question of whether all words encode meanings in the same way. I start from the assumption that they do not (*cf.* Blakemore 1987, 2002) and claim that aspects of the describing/indicating distinction are reflected in the fact that some words encode information that does not translate into the constituents of thought, but rather results in the *non-translational* activation of concepts, via *non-translational coding*.

---

[11] In her recent book, Robyn Carston wonders whether many *conceptual* encodings are 'not really full-fledged concepts, but rather concept schemas, or pointers to a conceptual space' (2002, p. 360). In particular she considers the word 'happy' and asks: 'Could it be that the word 'happy' does not encode a concept, but rather points to a conceptual region?'. This is an interesting point, and it would have clear implications for the distinction I am about to propose between translational and non-translational encoding. As Carston also points out, however, even if it turns out to be correct that some words encode 'pro-concepts', it will only be *some* words, and not all: 'There is a strong intuition that 'cat' encodes a concept CAT, which features in thought, and not just an abstract schema for constructing CAT* concepts or some pointer to knowledge about cats.' (ibid. , p. 362). In which case, the above distinction between translational and non-translational encoding is still motivated.

The essence of this distinction between translational and non-translational coding can be demonstrated using an analogy. Consider the following. There are two ways a friend might help you get from *A* to *B*. He might choose to take you in his car and drop you there directly, or he might simply point you in roughly the right direction, trusting that you will find your own way. If the destination represents a communicator's intended interpretation, this analogy reflects (albeit in highly intuitive terms) the difference between translational and non-translational coding (which corresponds very roughly to the speech-act distinction between describing and indicating and to the relevance-theoretic distinction between conceptual and procedural encoding (see below)).

With this analogy in mind, though, it is important to recognise that the distinction being made is one between two types of *coding*. A still further way of pointing someone in the direction of your intended interpretation is to provide evidence which involves no element of coding. Recall the examples from Chapter One where Lily deliberately and openly shows Jack her shiver, intending to communicate that she feels cold; or consider another in which I point at a cloud, intending to communicate that it's going to rain. Both of these cases, it could be argued, *result* in the non-translational activation of concepts. In both cases, however, the audience works out the communicator's intended interpretation in the absence of *any* code. The kind of non-translational activation I want to consider here is different in that it does contain a coded element that points the hearer in the appropriate direction, a direction they would not reliably take unless they knew the code.

As an example, consider (12):

(12) Jack: (referring to Lily) She's arrived.

Jack has clearly encoded something by uttering the pronoun 'she', but what? Wilson and Sperber (1993) argue that, looked at from a cognitive point of view, David Kaplan's (1977/1989) analysis of pronouns in terms of *character* and *content* amounts to the claim that, rather than directly encoding a conceptual representation of the intended referent,[12] what pronouns actually encode is a *constraint* which helps the hearer identify the intended referent in a given context by making a certain class of candidate referents more salient. Drawing on my earlier analogy, there are a variety of ways of 'pointing' someone in the appropriate direction. If your friend really *does* want you to get from A to B—and we invariably *do* want our interlocutors to infer the interpretation of our acts of ostensive communication—he might, in addition to pointing, tell you that B is a house with a flat roof.

In fact, the comparison with pointing is apt, for this is one of the central ideas behind the speech-act distinction between describing and indicating. Construed in terms of an inferential model, linguistic indicators do not determine a unique interpretation, but rather narrow the range of possible hypotheses from which the hearer must choose. Linguistic indicators are coded signals, but the code is non-translational: their function is to guide and constrain inference.

The view I am proposing is essentially one way of cashing out a distinction that has been explored in relevance-theoretic semantics between words that encode *concepts*—those words that result in translational activation, and *procedures*—those words that result in non-translational activation. The

[12] To propose that they did would be to suggest that pronouns are multiply ambiguous.

distinction was first presented by Diane Blakemore (1987) and taken further in Blakemore (2002), and the possibility of such a distinction is yet another consequence of adopting a fully inferential approach to ostensive communication.

In Blakemore's account, most words may be seen as encoding concepts, constituents of conceptual representations. Most of these contribute to the truth conditions of an utterance; they have logical properties, can act as input to inference rules, and are used to *describe* the world. Some words, however, do not map onto concepts. Rather than encoding constituents of conceptual representations, the function of these words is to constrain the inferential processes involved in constructing or manipulating these representations during the search for relevance. In effect, they guide the comprehension process by narrowing the hearer's search space and *indicating* the general direction in which the intended meaning is to be sought. There are a vast number of possible cognitive effects the speaker might have had in mind, and since processing effort is a factor in achieving relevance, such expressions will contribute to relevance by reducing the hearer's effort in finding the intended effects.

Consider Blakemore's analysis of the words 'so' and 'after all' in examples (13a-c):

(13a) Jack visits the dentist every six months. His teeth are good.

(13b) Jack visits the dentist every six months; *so* his teeth are good.

(13c) Jack visits the dentist every six months; *after all*, his teeth are good.

On Blakemore's account, in (13b) the word 'so' encodes a procedure which leads the hearer to process the two propositions in such a way that the first is a premise from which the second follows as a conclusion. In (13c) the expression 'after all' encodes a procedure which leads to the second proposition being understood as evidence for the first.[13] Blakemore's analysis classified them as examples of procedural expressions constraining inference at an implicit level (i.e. at the level of implicatures).

All this is not to suggest, however, that non-truth-conditional meaning is necessarily procedural. Consider example (14):

(14) Regrettably, your wisdom tooth will have to be extracted.

Despite the fact that 'regrettably' in (14) is non-truth-conditional (in the sense that it is not normally seen as contributing to the truth conditions of (14)), there are reasons to think that it does encode something conceptual (see Ifantidou-Trouki 1993, 2000). Firstly, it has conceptual counterparts which *do* contribute to the truth-conditions of utterances containing them, as in (15ab):

(15a) The incident at the dentist's was *regrettable.*

(15b) The dentist *regrets* her actions.

[13] Sperber (2001) proposes that words indicating inferential relationships (e.g. 'since', 'but' and 'nevertheless') might have evolved as 'tools of persuasion' in the cognitive arms race sparked by communicators' needs to display sufficient coherence in their argument that a hearer will accept their message. All such words have been analysed at various times as encoding *procedures*, and there may be some implications here for the evolution of procedural expressions. I return to evolutionary issues in Chapter Six.

Secondly, illocutionary adverbials such as 'frankly', which are not standardly seen as contributing to truth-conditions in (16a), combine compositionally with other expressions to form complex adverbial phrases, as in (16b):

(16a) Frankly, she's an absolute menace.

(16b) To put it frankly, and more frankly than I would dare if she had her drill in my mouth, she's an absolute menace.

This compositionality is to be expected if these adverbials encode conceptual representations, but it is hard to explain on a procedural account. This suggests an important modification to speech-act analyses of illocutionary force indicators, in that not all non-truth-conditional 'indicators' seem to work in the same way. The conceptual-procedural distinction thus cross-cuts the describing-indicating distinction (Wilson and Sperber 1993, Blakemore 2002).

What exactly does procedural information look like? Drawing on the distinction made in cognitive science between the representational and computational aspects of cognition, we might characterise it as providing computational *instructions* to the hearer: this is how it is often described in discussions of discourse connectives such as 'although', 'however', 'so', 'after all' following Blakemore 1987: Blakemore (1992, pp. 150-151) writes: '*But*, *after all*, *moreover* and inferential *so* do not contribute to a propositional representation, but simply encode instructions for processing propositional representations'. With other non-truth-conditional expressions, however, it might be better to view procedural content in a broader sense, as simply *activating* (or adding an extra layer of activation to) certain types of representations, or

contextual assumptions, or expectations about cognitive effects. Thus, a pronoun might activate a certain class of candidate referents from which the hearer must choose. We might characterise the procedural information encoded by mood indicators in this broad sense, as activating certain propositional-attitude descriptions used in the construction of higher-level explicatures, which the hearer is expected to draw on during the comprehension process. So Jack's utterance to Lily of (17a), may lead to her to form the higher-level explicature in (17b):[14]

(17a) Have you been to the dentist?

(17b) Jack is asking Lily whether she's been to the dentist.

We might, in fact, adopt the broader view for discourse connectives too. For what discourse connectives, mood indicators and pronouns have in common is that rather than translating into the constituents of conceptual representations[15] they involve non-translational activation. What is actually activated may be inferential rules or procedures, or contextual assumptions, or simply expectations of particular types of cognitive effects. In each case, the function of the non-truth-conditional expression is to guide the comprehension process by reducing the search space the inferential processes are working in and indicate—in Sperber and Wilson's words—'a rather abstract property of the speaker's informative

[14] Crucial to the relevance theory analysis of mood indicators is that the speech acts of saying, telling and asking are performed by utterances, not sentences. Interrogative mood is not necessarily linked with requests for information (see Wilson and Sperber 1988).

[15] In the case of the pronouns, the *output* of the procedure does provide a constituent. The constituent itself, however, is not encoded in the linguistic meaning of the pronoun.

intention: the direction in which the relevance of the utterance is to be sought' (1986/1995, p. 254).

In the next chapter I draw together various of the observations made in this chapter—in particular those concerning the showing-meaning$_{NN}$ continuum and the translational/non-translational, or conceptual/procedural distinction—and use them to analyse a group of 'semi'-natural expressions that have received much attention in the literature: interjections.

Chapter Four

> 'I should explain to you, Socrates, that our friend Cratylus has been arguing about names. He says that they are natural and not conventional—not a portion of the human voice which men agree to use—but that there is a truth or correctness, which is the same for Hellenes as for barbarians.'
>
> Hermogenes in Plato's *Cratylus*[1]

1. Interjections

Interjections are often regarded as marginal to language; while we feel them to be partly natural, we also feel them to be only partly coded (or conventionalised). Interjections seem to lie somewhere *between* showing and saying or meaning. This marginal linguistic status is reflected in various historical analyses. Latin grammarians described them as non-words, independent of syntax, signifying only feelings or states of mind. Nineteenth-century linguists regarded them as non-linguistic, or at best paralinguistic phenomena: 'between interjection and word there is a chasm wide enough to allow us to say that interjection is the negation of language' (Benfey 1869, p. 295); 'language begins where interjections end' (Muller 1862, p. 366). Sapir also described interjections as 'never more, at best, than a decorative edging to the ample, complex fabric [of language]' (1970, p. 7).

According to various definitions in the literature, 'interjections' represent a fairly heterogeneous class of items. Examples in English include *wow*, *yuk*, *aha*, *ouch*, *oops*, *ah*, *oh*, *er*, *huh*, *eh*, *tut-tut* (*tsk-tsk*), *brrr*, *shh*, *ahem*, *psst*, and even, according to some, *bother*, *damn*, (*bloody*) *hell*, *shit* (etc.), *goodbye*, *yes*, *no*, *thanks*,

[1] The source for this quote is *Plato: Collected Dialogues*, edited by Edith Hamilton and Huntington Cairns (see bibliography).

well. I will assume for the sake of argument that many of the above items *do* form a class, but will end up suggesting that interjections are very disparate and should not all be treated as contributing to communication in the same way.

Existing studies of the semantics and pragmatics of interjections raise three main questions:

(1) What do interjections communicate?

(2) How do interjections communicate?

(3) Are interjections part of language?

These questions have been approached from two largely opposite viewpoints. Ameka (1992), Wierzbicka (1992) and Wilkins (1992) argue that interjections are 'semantically rich and have a definite conceptual structure which can be explicated' (Wilkins 1992, p. 120). They treat interjections as part of language, and propose complex semantic analyses; I refer to this as the *conceptualist* view. Others, notably Goffman (1981), contend that an interjection 'doesn't seem to be a statement in the linguistic sense'. Rather, it is 'a ritualised act, in something like the ethological sense of that term' (1981, p. 100). Interjections, according to this view, are not part of language, and are analysed in terms of the socio-communicative roles they play, rather than any linguistic content they may have.

In the light of the above questions, the aim of this chapter is to assess the relative strengths and weaknesses of these two contrasting approaches and to suggest a new analysis of interjections which preserves the insights of both. This analysis will build on some of the discussion in the previous chapter. In particular, it will also make use of (a version of) the showing-meaning$_{NN}$

continuum and also the relevance theoretic distinction between translational and non-translational (or conceptual and procedural) encoding. The conceptual-procedural distinction will be further discussed in Chapter Five, for the question remains whether, given the marginal linguistic status of interjections, an analysis with its roots in what is essentially a *linguistic* distinction is appropriate.

The view that interjections have, at best, marginal linguistic status can still be found in the contemporary literature: Quirk, Greenbaum *et al* (1985, p. 853) describe interjections as 'purely emotive words which do not enter into syntactic relations'; Trask (1993, p. 144) describes an interjection as 'a lexical item or phrase which serves to express emotion and which typically fails to enter into any syntactic structures at all'; Crystal (1995, p. 207) concurs—'an interjection is a word or sound thrown into a sentence to express some feeling of the mind'.

There are exceptions, though. As noted above, conceptualists see interjections as properly linguistic, with rich semantic structures. However, whilst the conceptualists are agreed that since they have semantic structure, interjections are part of language, they do not agree on what exactly an interjection is. Introducing the conceptualist view, Ameka (1992) divides interjections into two main classes: *primary* and *secondary* interjections. Primary interjections are words that cannot be used in any other sense than as an interjection, e.g. *oops* and *ouch* in (4):

(4) Patient: Be careful with that needle!

Dentist: *Oops*.

Patient: *Ouch*!

These items are non-productive in the sense that they do not inflect and are not movable between word-classes. Secondary interjections 'are those words which have an independent semantic value but which can be used...as utterances by themselves to express a mental attitude or state' (Ameka 1992, p. 111), e.g. *hell* and *shit* in (5):

(5) Dentist: *Hell*! I'm sorry.

Patient: *Shit*! Get the bloody thing out of my cheek!

Both types of interjection are syntactically independent, in that they can constitute an utterance by themselves, and are only loosely integrated into the grammar of the clause containing them. When written, interjections are separated off from the main clause by means of a comma or exclamation mark. Furthermore, Ameka observes, they 'always constitute an intonation unit by themselves' (1992, p. 108).

Wierzbicka's definition of an interjection correlates closely with Ameka's conception of a primary interjection. She suggests that it is preferable not to regard exclamations such as *shit* and *hell* as interjections, since their semantics should be included in the semantics of the nouns/verbs they are derived from: I shall follow her on this. While Ameka's definition is too broad for her, for Wilkins it is too narrow. He uses a variety of hedges in his formal definition of interjections (1992, p. 124), which 'catches elements that would be called "secondary interjections" ... "interjectional phrases" and "complex interjections" by Ameka' (1992, p. 125). There is thus no general agreement on how interjections can be defined.

Since Goffman (1981) does not regard interjections as part of language, he does not define them in the same way. In fact, for the majority of expressions I shall look at in this paper, he prefers the term *response cry*: 'We see such "expressions" as a natural overflowing, a flooding up of previously contained feeling, a bursting of normal restraints' (1981, p. 99). By 'response cry', Goffman is referring primarily to expressions such as *ouch*, *oops*, *yuk*, *wow*, *eh*, *ah*, *aha*, *oh* etc., which he regards as non-words. Since 'nonwords as a class are not productive in the linguistic sense, their role as interjections being one of the few that have evolved for them...[they] can't quite be called part of language' (1981, p. 115). However, he does grant that since these cries are found cross-linguistically, and since certain forms stabilise within a given speech community, the term *semiword* might be appropriate. Swear words are of course highly productive. But while conceding that they are probably more a part of language than non-words such as *oops* and *ouch*, he does not see this as reason to exclude them from the class of response cries, which in his view exist on a continuum between display and properly linguistic items.

One point of agreement between the conceptualists and Goffman is that *an interjection is capable of constituting an utterance by itself in a unique, non-elliptical manner.* Another point accepted by both camps is that interjections are tied to emotional or mental attitudes or states. From the examples on my introductory list, *wow* might be said to express excitement, delight, wonder etc., *yuk* to express disgust or revulsion, *ouch* pain, *aha* surprise etc. Wierzbicka suggests that alongside these *emotive* and *cognitive* interjections, there are some *volitive* ones, used to express wants or desires: *psst*, *ahem*, *shh* and *eh*, for example, serve as requests for attention, quiet or confirmation. A second

criterion, then, by which we might classify an expression as an interjection is that *an interjection expresses a mental or emotional attitude or state.*

These two criteria seem to me to form an adequate working characterisation. In what follows, I will retain the conceptualists' primary/secondary distinction, and focus mainly on primary interjections, which have no counterparts in other syntactic categories. Focussing on primary interjections also allows me to largely abstract away from linguistic expressions such as *yes*, *no*, *thanks* and *goodbye*, which could be seen as fitting the above criteria, but are not central to the claims of this chapter. I will, however, consider the status of certain stylised imitations, such as 'ha ha', 'boo hoo' etc.

2. Interjections and concepts

'interjections have real 'semantic' (i.e. propositional/conceptual) content...'
(David Wilkins 1992, p. 119)

The conceptualists would presumably answer questions (1), (2) and (3) along the following lines: first, interjections communicate complex conceptual structures; second, communication is achieved principally by means of *encoding* conceptual structures; third, since interjections are viewed as having 'semantic' content, they are part of language.[2] Below in (6) is an example of the kind of analysis the

[2] Of course, a great deal depends on how you interpret the word 'semantic' here. Wierzbicka (2000), for example, discusses the 'semantics' of human facial expression, which suggests she has a somewhat broader conception of the notion than the one adopted in this thesis, where 'semantics' is taken to be the study of *linguistic* meaning. Despite these terminological differences, however, I think that on the strength of the quotes from Ameka and Wierzbicka below (p. 165 and p. 167 respectively), I am justified in taking it that according to the conceptualist view interjections are part of language. See the next chapter for discussion of Wierzbicka's conceptualist analysis of facial expressions.

conceptualists propose, Wierzbicka's conceptual structure for *wow* (1992, p. 164):

(6) *wow!*

I now know something

I wouldn't have thought I would know it

I think: it is very good

(I wouldn't have thought it could be like that)

I feel something because of that

As can be seen from this analysis, conceptualist analyses of interjections are massively decompositional, and should be viewed in the wider context of Wierzbicka's programme to develop a Natural Semantic Metalanguage. This approach is based on a set of around fifty primitives, designed to represent the innate building blocks of meaning: 'research of recent years has proved Wittgenstein wrong... words can be rigorously defined' (Wierzbicka 1994, p. 433). Wierzbicka extends this approach to interjections: 'we can capture the subtlest shades of meaning encoded in interjections relying exclusively on universal or near-universal concepts such as "good" and "bad", "do" and "happen", "want", "know", "say", or "think"' (Wierzbicka 1992, p. 163).

Although many subtle and intuitively appealing analyses have been proposed within this framework, there are several problems with the approach. Firstly, there are serious objections to decompositional accounts of meaning. Fodor, Fodor and Garrett (1975) provide experimental psycholinguistic evidence against decompositions containing negative elements. If the concept [bachelor]

decomposes into a complex containing a negative concept [unmarried], then difficulties associated with processing and evaluating the validity of arguments containing interactions between quantifiers and negative items should arise processing arguments containing the word 'bachelor'. However, in tests, this was found not to be the case; the argument in (7a) is easier to process and evaluate than (7b), even though on a decompositional approach they are semantically equivalent:

(7a) If practically all the men in the room are bachelors, then few of the men in the room have wives.

(7b) If practically all the men in the room are unmarried, then few of the men in the room have wives.

This objection applies directly to Wierzbicka's analysis of the interjection in (6), which also contains negative elements. While there are obvious problems applying the above test to interjections, which do not integrate into syntactic structure, and to definitions such as (6), which are too long to be satisfactorily embedded in their entirety, the proposal that *wow* encodes a negative element is not supported by the data in (8ab); (8a) is easier to process and evaluate than (8b), suggesting it does not contain a negative element:

(8a) If the fireworks were good and he didn't say *wow*, he wasn't really impressed.

(8b) If the fireworks were good and he didn't say he wouldn't have thought he would know it, he wasn't really impressed.

Wierzbicka's structures for *oops* (1992, p. 163) and *yuk* (1992, p. 168) also contain negative elements, as do Wilkins' for *ow* (1992, p. 149) and *wow* (see (10) below), and the same objection applies to these structures too.

Fodor (1981) provides further arguments against decompositionalism. Very few words, he claims, are decomposable into satisfactory definitions: in this respect, the classic example 'bachelor' is exceptional. Fodor demonstrates that the task of analysing other relatively simple words into necessary and sufficient conditions is a hopeless one.[3] He takes the word 'paint' as an example, and argues that *x paints y* is not satisfactorily defined as *x covers y with paint.*[4] To support his claim, he raises a series of objections, each of which he attempts to counter with a more complex definition. When an explosion at a paint factory covers a passer-by with paint, the factory has not painted the passer-by: perhaps, then, the definition should stipulate an *agent*. However, in covering the surface of the ceiling of the Sistine Chapel, Michelangelo, while certainly an agent, was not painting the ceiling, but rather painting *a picture on* the ceiling. With these counter-examples in mind, Fodor defines *x paints y* as meaning *x is an agent and x covers the surface of y with paint, and x's primary intention in covering the surface of y with paint was that the surface of y should be covered with paint in consequence of x's having so acted upon it.* However, he finds a counter-example to even this most complex definition. For when Michelangelo dipped his brush in his paint pot, the above conditions were satisfied, but he was not painting his

[3] See Wierzbicka (1996, pp. 253-257) for her response to Fodor.

[4] Fodor's discussion (pp. 287-290) is based on a definition originally presented in Miller (1978, p. 285).

paintbrush: 'when it comes to definitions', Fodor concludes, 'the examples almost always don't work' (1981, p. 288).[5]

Along similar lines we can find counter-examples to the conceptualist structures for interjections. Firstly, the definition in (6) includes the line *I think: it is very good.* But this overlooks the fact that *wow* can just as easily express negative feelings, such as outrage, or disgust;

(9) *Wow*! That's outrageous!

Wow! That's disgusting!

This point is also is raised by Wilkins (1992, p. 150). To account for it, and the fact that neither Wierzbicka's nor Ameka's definition captures the immediacy of the kind of reaction expressed by an utterance of *wow*, he proposes the more complex structure below (1992, p. 151):

[5] Fodor maintains this view in Chapter 3 of his 1998 book: *Concepts: Where Cognitive Science Went Wrong*. 'There are practically no defensible examples of definitions; for all the examples we've got, practically all words (/concepts) are undefinable. And of course, if a word (/concept) doesn't have a definition, then its definition can't be its meaning.' (1998, p. 45)

(10) "*wow!*"

I_U have just now_T become aware of $this_I$ something,

that I_U wouldn't have expected

[or 'that I_U wouldn't have thought I_U would become aware of']

$This_I$ something is much more $X_{[Pr\text{-}of\text{-}this\ I]}$ than I would have expected,

and this causes me_U to feel surprised,

and to feel that I_U could not imagine this something being more $X_{[Pr\text{-}of\text{-}this\ I]}$

than it already is now_T.

I_U say '/wau!/' because I_U want to show how surprised (and impressed)

I_U am feeling right now_T.

But as with Fodor's more complex definitions, there are still problems. For example, there are aspects of the meaning of *wow* that the structure in (10) does not adequately capture. Does '*this is much more X than I would have expected and...causes me to feel surprised*' 'rigorously' define the subtle shades of positive meaning that an utterance of *wow* might communicate? From surprise and being mildly impressed, through amazement and astonishment to jaw-dropping bewilderment? From satisfaction through enjoyment to absolute exhilaration? Also, is it true that *wow* communicates that the speaker feels they '*could not imagine this something being more X than it already is*'? Does a spectator at a firework display communicate that he feels that this is the most spectacular firework he can imagine when he utters *wow*? Fodor's point that there are always counter-examples to be found, no matter how complex the definition, appears to hold for interjections too.

The second problem with the conceptualist approach is that an utterance of *wow* seems to communicate something altogether vaguer than the kind of structures they propose would predict: as suggested above, the meaning of *wow* surely cannot be 'rigorously defined'. This is not to deny that interjections can communicate a great deal. However, the range of communicative effects an utterance of *wow* might give rise to, when combined with different intonations and facial expressions, seems to go well beyond anything capturable in a small set of conceptual structures such as those proposed above.

An analogy with some of the other natural behaviours humans use to communicate is instructive here. What a speaker might communicate by using an affective tone of voice seems too nebulous to be paraphrased by a fixed structure such as (6). A facial expression or gesture might convey more than a string of words ever could, but it is not obvious that it is *encoding* anything.

The context-dependence of interjections is the third problem for the conceptualist approach. Of course, (6) is not a fully propositional structure, because it contains uninterpreted indexicals (*I*, *it*, *now*) which are assigned reference by means other than linguistic decoding. Wilkins employs a variety of deictic sub-scripts (see (10)) to account for this context-dependency: 'each deictic element must be filled referentially before the interjection can be fully meaningful' (1992, p. 137). But the communicative content of interjections is *so* context-dependent that it seems implausible to suggest that the only contribution of pragmatic/contextual factors to their interpretation is the assignment of reference to indexicals. The conceptualist approach fundamentally underestimates the contribution of pragmatic/contextual or *inferential* factors to the interpretation of interjections. I will return to this point below.

The vagueness and context-dependence of interjections also relate to a fourth, more general problem with the conceptualist account. As mentioned above, humans use a wide range of behaviours to communicate. Consider, for example, how an individual might convey a feeling of pain. Methods range from allowing someone to see an entirely natural and instinctive contorted facial expression, to a scream such as 'aaaargh', to a culture/language-specific *ouch*, to a fully linguistic 'it hurts like hell'. No one would propose that grimaces or screams encode conceptual structure, but communicate they do. Interjections retain an element of naturalness and spontaneity that suggests they fall somewhere *between* the natural and the linguistic. With tone of voice, facial expressions and even gestures, they share the property of being partly natural: the conceptualist approach overlooks this.

A fifth problem is that intuitions do not support the claim that interjections encode the kind of conceptual structure the conceptualists propose. Consider (11) below, Wilkins' conceptual structure for *ow* (Wilkins 1992, p. 149):

(11) *"ow!"*

I suddenly feel a pain (in this part of my body) right now that I wouldn't have expected to feel.

I say '[au!]' because I want to show that I am feeling pain right now [and because I know that this is how speakers of English can show (other speakers of English) that they are in pain (in a situation like the situation here)]

While one is happy to concede that the italicised expressions in (12ab) express the same (or similar) concepts, it is not obvious that the same is true of those in (13ab), which do not *feel* synonymous in the same way:

(12a) Be careful with that *needle*!

(12b) Be careful with that *hypodermic*!

(13a) *Ow*! What did you do that for?

(13b) *I suddenly feel a pain* etc. What did you do that for?

It could, of course, be our unfamiliarity with the sheer complexity of the conceptual structure in (11) that is responsible for this intuition. However, even if we strip the conceptual structure down to its bare essentials, where *ow* encodes something like 'I feel pain', there are still problems. (14a), for example, intuitively involves a conceptual repetition, while (14b) does not:

(14a) I feel pain, I feel pain.

(14b) *Ow*, I feel pain.

And interjections are not interchangeable with their conceptual counterparts; they do not, for example, occur in embedded positions:[6]

[6] An anonymous referee of a published version of this chapter (see Wharton 2003a) points out that the non-embeddability of interjections (and 'expressive elements and constructions' generally) is also central to Banfield's (1982) account of represented speech and thought (essentially, *style indirect libre*). (Although it should be noted that Banfield is concerned with constraints on embedding in 'that' clauses (1982, p. 30-32).)

(15a) If I feel pain, I'll tell you.

(15b) * If *ow*, I'll tell you.

In recent unpublished work, the philosopher David Kaplan (1997) has addressed (among other things) the linguistic difference between 'I feel pain' and *ouch*. Better known for his work on indexicals, Kaplan sees similarities between indexicals on the one hand, and expressives (interjections—*ouch*, *oops*) and epithets ('the bastard') on the other: all these expressions, he claims, are better analysed in terms of a *Semantics of Use* rather than (or as well as) a *Semantics of Meaning*. To account for the difference between 'I feel pain' and *ouch*, he introduces the notions of *descriptive* and *expressive* content: while 'I feel pain' has descriptive (truth-conditional/propositional) content, *ouch* has expressive (non-truth-conditional/non-propositional) content. This distinction is similar to the distinction drawn by speech-act theorists between describing and indicating mentioned in Chapters Two and Three; I return to this below.

Kaplan's notion of descriptive content does seem to parallel the conceptualists' notion of conceptual/propositional content. In this case, the descriptive/expressive distinction supports the above intuitions that one of the reasons *ow* and 'I feel pain' are not interchangeable in (14ab) and (15ab) is that *ow* does not encode conceptual structure. In Kaplan's terms, the *modes* of expression are different.

The sixth problem relates to the fact that interjections do not contribute to the truth conditions of the utterances that contain them. In fact, the non-truth-conditionality of interjections may be one of the factors responsible for the intuitions in (14ab). Consider (16ab):

(16a) I feel pain, the anaesthetic isn't working.

(16b) *Ouch*, the anaesthetic isn't working.

(16a) makes two assertions: it is true when and only when the speaker feels pain and the anaesthetic isn't working; (16b) only makes a single assertion, and is true if and only if the anaesthetic isn't working. The dentist could not respond to a patient's utterance of '*Ouch*!' in (4) with: 'You're *lying*, you can't feel any pain'. As noted above, conceptual representations have logical properties, and are capable of being true or false. As a result, a conceptual representation can contradict or imply other conceptual representations and act as input to logical inference rules. Since interjections do not seem to have these properties, it might be best to treat them as not encoding fully conceptual structures.[7]

To summarise, there are six problems with the conceptualist approach: firstly, there are problems with decompositionalist accounts of meaning generally; secondly, the communicative content of interjections is vaguer than the proposed conceptual structures would predict; thirdly, the highly context-dependent nature of interjections suggests a substantial pragmatic contribution to their comprehension; fourthly, the approach overlooks the fact that interjections share with certain paralinguistic behaviours the property of being partly natural; fifthly, the fact that they do not appear to be synonymous with their fully conceptual counterparts suggests they do not encode concepts; sixthly, the non-truth-conditional status of interjections suggests that a conceptual account is inappropriate, and that alternative semantic treatments should be explored.

[7] This is not to say that all conceptual meaning is truth-conditional meaning (see example (14) and (16a) in Chapter Three). This issue is addressed more directly in Section 5 below. For further discussion see Wilson and Sperber 1993. Ifantidou 2001.

3. Interjections and 'response cries'

> 'During the Wimbledon tennis championships in 1981, officials were confronted with an unusual problem. Some male players, notably Jimmy Connors, were regularly grunting loudly as they hit the ball. Their opponents... claimed the noises were distracting and were emitted deliberately to throw off their timing. When officials confronted Connors... he explained that he had no control over his grunting; it just happened when he hit the ball hard... Wimbledon officials then observed the different players, trying to discern which grunts were intentional and which were not.'
>
> (Seyfarth and Cheney 1992, p. 78)

Goffman (1981) discusses interjections in terms of the socio-communicative roles they play rather than any linguistic content they may have. Of the questions that are the focus of this chapter, he is concerned with questions (1) and (3), and not question (2).

He considers three types of 'roguish utterances', which violate the conditions that normal 'talk' observes: *self-talk*, *imprecations* (swearing) and *response cries*. It is the latter two which are relevant here, and Goffman's distinction between response cries such as *oops*, *ouch*, *wow* etc. and imprecations reflects the conceptualists' primary/secondary interjection distinction discussed in the last section.

Goffman would not support Jimmy Connors' claim that his grunts were unintentional. Indeed, his primary concern is the fact that such sounds are invariably intended for the benefit of others. The purpose of *strain grunts*, for example, is often to warn others to stand clear. He comments, 'these sounds are felt to be entirely unintentional, even though the glottis must be partially closed off to produce them and presumably could be fully opened or closed to avoid doing so' (1981, p. 105): Goffman fifteen, Connors love.

Goffman classifies response cries according to the function they serve. Some indeed exploit more or less instinctive, natural reactions: the *transition display*,

where a person uttering *brrr* when leaving a warm atmosphere for a cold one might not only do so to restore some sort of physical equilibrium but also to 'fall into cadence with the others in the room' (1981, p. 101); the *spill cry*, where a person uttering *oops* on dropping something might do so because it has the effect of 'downplaying import and hence implication as evidence of our incompetence' (1981, p. 102). According to Goffman, the main function of *ouch* (the *pain cry*) is to warn others that a threshold for pain is being reached, or about to be breached. Such response cries are not productive linguistically and are therefore peripheral to language proper.

Imprecations, by contrast, are highly productive linguistically. However, Goffman notes that an exclamation of *shit!* 'need no more elide a sentence than need a laugh, groan, sob, snicker or giggle—all vocalisations that frequently occur except in the utterances ordinarily presented for analysis by linguists'. Nor does it help 'to define *shit!* as a well-formed sentence with *NP!* as its structure'. He concludes that 'imprecations, then, might best be considered...as a type of response cry' (1981, p. 112).

One of the most important points that Goffman raises is that there may be a continuum of elements between the properly linguistic and the non-linguistic, or between display (or showing) and saying. Since *ouch*, *oops* etc. are not productive linguistically, according to Goffman they 'can't quite be called part of language' (1981, p. 115). Because of their productivity, imprecations are part of language (17abc) (though recall that when used as interjections they are non-productive):

(17a) That dentist is shit.

(17b) The dentist got really shitty with me.

(17c) He was the shittiest dentist I've ever had the misfortune to see.

The distinction, however, is not clear-cut: 'response cries such as *eek!* might be seen as peripheral to the linguist's domain...but imprecations...are more germane, passing beyond semiword segregates to the traditional material of linguistic analysis' (1981, p. 121).

An illustration of Goffman's proposal might be as follows: to show someone you are delighted with a gift you allow them to see your natural reaction, a smile; to tell them you are delighted you utter something like 'it's wonderful!'; to utter an interjection like *wow* is to communicate that you are delighted by adding a certain element of *coding* which takes it beyond mere display, but falls short of language proper.

There are clear parallels between Goffman's continuum and the continuum sketched in the last chapter. Both deal with the whole range of communicative phenomena: from 'natural' display to the fully linguistic. Indeed, one of the aims here (and in my final chapter) is to examine ways in which this type of continuum, which—as we shall see—other theorists have proposed, might mesh with the continuum presented in Chapter Three. However, it's worth pausing to notice that in contrast with the continuum presented in the previous chapter, Goffman's continuum appears to be rooted in the role coding—and coding alone—plays in communication. It's unclear to me precisely what Goffman means by 'display', but since there is no discussion of the role played in human communication by the expression and attribution of intentions in communication,

and since in other published works (see, for example, Goffman 1964) Goffman talks of the 'rituals' and 'regulations' that 'govern' conversational exchanges, I am taking his continuum to be one between non-linguistic and linguistic coding.

Goffman, as we have seen, does regard it as important that some response cries are 'intentional', but he does not appear to be using the word in the 'rich, philosophical sense' mentioned in my introduction. I return to these issues in Chapter Six, and argue that while a continuum of the kind outlined in the previous chapter can accommodate a continuum of the kind envisaged by Goffman (and others), the reverse is not the case.[8]

Although he sees response cries as outside language proper, a strength of Goffman's account is that he is keen to illustrate their communicative adaptability. He points out that if you are being told by a friend about a particularly gruesome moment from their last trip to the dentist's, you might utter *ouch* sympathetically on their behalf.[9] Or it might be used as in (18):

(18) Dentist: That'll be £75 for the consultation and £30 for the cavity.
Patient: *Ouch!*

Here again Goffman is distancing himself from the view that primary interjections are a simple 'natural overflowing'. It is, after all, intuitively clear

[8] I hope I am not attributing to Goffman views he never held; as I say above, the difference between the two types of continuum will (hopefully) become clearer in the final chapter. Until then, I would like the reader to bear in mind that since both types of continuum deal with the whole range of communicative phenomena—from display to the fully (linguistically) coded—the similarities between them are worth holding onto.

[9] Both this use and Goffman's 'warning' example are 'pragmatically determined variants' according to Wilkins (1992, p.150*n*.). He says nothing of the use in (18).

that while they are instinctive in some respects, *ouch* and most primary interjections are under our conscious control. If I bring a hammer down forcefully on my thumb, the four-letter word I utter is unlikely to begin with 'o'.[10] A person screaming in agony does not scream *ouch!*. In this respect, we should be careful not to overestimate the expressive, instinctive nature of these primary interjections.

There are many interesting ideas in Goffman (1981). The question of what interjections communicate is, in almost all cases, beautifully explicated. In terms of the questions asked at the beginning of this chapter, the problem is that he says nothing about *how* interjections communicate. In this respect, whilst it affords some insights that are certainly worth preserving, his analysis does not provide a satisfactory theoretical alternative to the conceptualist approach. In the next section, I will look at some analyses of linguistic meaning which offer some alternatives to the conceptualist account of interjections.

4. Interjections and meaning: *what do interjections communicate*?

In Chapters Two and Three we have seen that over the last 30 years, philosophers of language and linguists have explored the idea that not all linguistic meaning is descriptive, or conceptual. At various times a distinction has been made between *truth-conditional* and *non-truth-conditional*, or *propositional* and *illocutionary* content, and between *describing* and *indicating*, or *saying* and *conventionally implicating*. If interjections do not encode

[10] Though, as Goffman points out, it might if I were helping out at the local playgroup.

descriptive, or conceptual meaning, it is worth exploring whether they can be analysed as non-truth-conditional indicators of some kind.

Recall from Chapter Three that, in relevance-theoretic terms, the proposition expressed by an utterance is the *basic-level explicature*. This explicature (already the result of pragmatic inference—leading to disambiguation, reference resolution and various types of enrichment) also figures in various *higher-level explicatures*, the construction of which requires the embedding of the proposition expressed under a speech-act or propositional-attitude description. In this way, aspects of both speech-act theory and Gricean pragmatics are retained within the relevance theory framework.

To illustrate this approach once more, consider how Jack might interpret Lily's utterance in (20a). Having recovered the proposition expressed, he might embed it under a speech-act description, as in (20b), or a propositional-attitude description, as in (20c). These would be higher-level explicatures of Lily's utterance in (20a):

(20a) Lily (regretfully): I've got a toothache.

(20b) Lily is saying that she's got toothache.

(20c) Lily regrets that she's got toothache.

The framework as presented so far suggests a way we might approach question (1)—What do interjections communicate?. Interjections might be indicators of higher-level explicatures, containing the type of speech-act or propositional-attitude information the hearer is expected to infer. A candidate for an interjection that might encode a similar sort of information to interrogative

mood indicators—*cf.* example (17a) from the previous chapter—(although it is not as integrated into the syntax) is *eh*. Thus, in relevance-theoretic terms, a patient interpreting the dentist's utterance in (21a) might form the higher-level explicature in (21b), or perhaps (21c):

(21a) Dentist: So you're having three teeth out, *eh*?

(21b) The dentist is asking whether I'm having three teeth out.

(21c) The dentist is requesting confirmation that I'm having three teeth out.

In many languages such particles appear to be fully grammaticalised. Japanese has an interrogative particle 'ka', added to the end of an interrogative utterance. Wilson and Sperber (1993) point out that certain dialects of French have an interrogative particle 'ti' which performs the function carried out by word-order in other dialects, and might be analysed along similar lines to 'eh'. Indeed, in English a similar questioning attitude toward the proposition is often conveyed by the word 'right?', or the tags 'aren't you?' or 'are you?'.

Wilson and Sperber (1993) also propose that the English interjection *huh* might be used to encourage the construction of higher-level explicatures involving a dissociative attitude toward an attributed utterance or thought. Consider (22a), which might lead a hearer to derive the higher-level explicature in (22b):

(22a) Lily: Dentists are human, *huh*!

(22b) It's ridiculous to think that dentists are human.

Cross-linguistic data suggest that many languages contain particles that might be analysed in a similar way. Japanese (Itani 1995) and Sissala (Blass 1990) have hearsay particles, [ttɛ] and [rɛ] respectively, which mark propositions as attributed to another speaker (or thinker). Sadock and Zwicky (1985, p. 161) note that Lahu has 'a very large number of particles that indicate attitudes, rational or emotional, toward a proposition'.

Since a feature of interjections in general is that they express attitudes, we might consider the extent to which these attitudes are similar to those conveyed in example (20a). In a similar way, utterances of (23a) and (23a) might lead a hearer to form the higher-level explicatures in (24b) and (24b):

(23a) *Aha*! You're here.

(23b) The speaker is surprised that I am here.[11]

(24a) *Wow*! You're here.

(24b) The speaker is delighted that I am here.

In speech-act terms both *aha* and *wow* in (23) and (24) can be analysed as performing expressive speech acts. In fact, all the examples I have considered so far seem to fit the speech-act framework, in that there appears to be an attitude, emotional or otherwise, being conveyed toward the proposition expressed—satisfying John Searle's (1979) definition of an *expressive* speech-act: 'the illocutionary point of this class is to express the psychological state specified in

[11] I address the question of whether the hearer might or might not form the further higher-level explicature 'the speaker is *saying* that he is surprised I am here' in Section 6 (the same applies to (24ab)).

the sincerity condition about a state of affairs specified in the propositional content' (Searle 1979, p. 15).

Consider (25ab), however:

(25a) *Yuk*! This mouthwash is foul.

(25b) *Wow*! This ice cream is delicious.

Here, the attitudes being expressed are not being expressed to an embedded proposition. Utterances of these sentences would not lead a hearer to form the higher-level explicatures in either (26a') or (26b'):

(26a') The speaker is disgusted that the mouthwash is foul.[12]

(26b') The speaker is delighted that the ice cream is delicious.

It seems that in these examples, attitudes are being expressed to objects rather than propositions: in the case of *yuk*, to the mouthwash (or more particularly the taste of it), and in the case of *wow* to the ice cream (or the sight or taste of it). As another example, consider (27):

(27) Child: (taking foul-tasting medicine) *Yuk!*

Here, the interjection stands alone as an utterance in its own right in the unique non-elliptical manner characteristic of interjections. Not only is the attitude not directed at any embedded propositional content, there *is* no propositional content

[12] A dentist might chastise her sloppy assistant by saying 'I am disgusted that this mouthwash is foul', but would not communicate this by uttering (26a).

to embed. For this reason, it is hard to analyse (27) as conveying a higher-level explicature or expressive speech act since there is no linguistically encoded logical form to embed under it.

In fact, we might ask whether what is communicated by the interjections in (26ab) and (27) are properly described as emotional attitudes at all. In (27) in particular, what the interjection communicates seems to be something more like a 'feeling' or a 'sensation', at least according to some characterisations of those terms.

Rey (1980) characterises emotional states in terms of an interaction between several elements: *cognitive*, *qualitative* and *physiological*. Thus, sadness is characterised as involving an interaction between a cognitive element—the knowledge that something has happened which you would prefer not to have happened, or the belief that something which you would prefer not to happen is going to; a qualitative element—that feeling of being 'down' (perhaps accompanied by behaviour consistent with feeling this way, such as drooping shoulders and a flat tone of voice); and a physiological element—chemical changes in the brain (in the case of sadness or depression, depletion of norepinephrine). Whilst emotional states crucially involve cognitive, as well as qualitative and physiological elements, feelings or sensations need not. Seen in these terms, what is communicated by *yuk* in (27) is indeed a feeling or sensation rather than an emotion, and not an emotional attitude or propositional attitude proper. It seems, then, that the framework as being presented so far is too restrictive: perhaps it is not possible to account for the meaning of interjections *solely* in terms of propositions and propositional attitudes, as existing speech-act and relevance-theoretic analyses seem to suggest.

As well as the example in (27), other interjections, such as *ouch* (see (4)), are difficult to account for in terms of propositional attitudes; these might also be said to communicate feelings or sensations rather than emotions: the speaker simply reveals something about her internal state. In Kaplan's terms this state is *expressed* rather than *described.* In cognitive terms, we might cash this out by proposing that there is something *non-representational* about what interjections encode. This proposal would be consistent with the arguments presented in Section 2, and is one I explore in the next section.

The question of what interjections communicate, then, requires various answers. In some cases they might be analysable as conveying a speech-act or propositional-attitude information. In this regard, interjections such as *eh* and *huh* pattern with discourse particles such as those I mentioned earlier. The interjection *alas* also might express a propositional-attitude proper. Thus, instead of sighing regretfully, and speaking in a regretful tone of voice, Lily might preface her utterance with *alas*, and in doing so express her attitude of regret more explicitly.

Other interjections (e.g. those in (23a), (24a)) also express propositional attitudes: emotional attitudes expressed toward propositions in the sense suggested by Searle. However, in some instances what an interjection expresses might be directed toward a percept or object which is the cause of a qualitative or physiological response, and not to a proposition (e.g. (25ab)). In these cases, whether or not what is communicated is an *emotional* attitude depend on whether there is a cognitive element interacting with the qualitative and the physiological. The cognitive element is not always present: in fact, it could be argued that interjections are *primarily* directed at the percepts and objects that are the causes

of particular responses, and only by extension at propositions. Finally, some uses of interjections (see (27), (4)) clearly communicate feelings or sensations, and not propositional attitudes proper.

An adequate analysis of what interjections communicate should take account of all these observations. It should also address the fact that whatever interjections communicate—propositional attitudes, emotions, feelings or sensations—it does not seem to be achieved by encoding conceptual representations. I turn to this question in the next section.

5. Interjections and procedures: *how do interjections communicate*?

In the previous chapter I introduced the distinction between translational and non-translational encoding, based on the relevance-theoretic distinction between conceptual and procedural encoding. Having argued against conceptualist (or translational) accounts of interjections, I now want to explore the possibility of a non-translational/procedural approach. Many of the arguments I will use are based on those developed to test whether words encode conceptual or procedural information (for that reason I favour the term 'procedural' rather than 'non-translational' in much of what follows). In the next chapter, I will consider whether an analysis with it roots in *linguistic* distinctions, is appropriate for items such as interjections, which have at best marginal linguistic status.

As we saw in Section 2, we already have good evidence against conceptual accounts generally. Furthermore, the tests described in the previous chapter for conceptual or procedural content seem to support a procedural account. Interjections have *no* synonymous truth-conditional counterparts; they are

linguistically *non*-productive and do not undergo compositional semantic rules. It seems plausible to suggest, then, that they encode procedural information which 'points' in the general direction in which relevance should be sought.

In an account developed along these lines the procedural information encoded in interjections might be seen as activating various attitudinal concepts or types of concepts, but do so along non-translational lines. Under such an account *wow* would not encode a concept that a hearer translates as 'X is delighted'. Instead it would activate (or add an extra layer of activation to) a range of attitudinal descriptions which involve delight, surprise, excitement etc. In the case of *yuk*, the attitude would be one of disgust; in the case of *aha* it would be one of surprise etc. In the case of *eh*, what would be activated is a range of interrogative propositional-attitudes descriptions; in the case of *huh*, it would be a range of dissociative attitudes, and so on. Intonation and facial expression might provide further clues to the particular attitude involved.

What a hearer does with the attitudinal or speech-act information activated might vary in different situations. In utterances of (23a) and (24a), a hearer might use it to construct a higher-level explicature. It is hard to seen how (25ab) might lead to propositional embedding, and it may be that many interjections are primarily directed at suddenly perceived objects and events, and only by extension to propositions.

This kind of account would square nicely with the observation made in the last section that there is something non-representational about what interjections encode. It also means that we might see some interjections as working in a similar manner to more fully grammaticalised discourse particles—'please', 'well', 'then', 'now'—with which they share a lack of syntactic integration.

It would also resolve five of the six problems with the conceptualist account outlined above: firstly, the approach is clearly non-decompositional; secondly, the result of comprehension may be vague, since a wide range of possible propositional-attitude descriptions may be equally activated, and there may be no way for the hearer to choose among them.

As to the third problem, the precise conceptual structure actually arrived at by the hearer will be different in different contexts, since the particular interpretation is the outcome of several overlapping inferential processes it constrains, rather than simply being decoded. Even in the case of *eh*, one of the best candidates for being linked to a particular speech-act—i.e. a request for confirmation—it would be unsatisfactory to propose that this is what is encoded. Consider (28) below:

(28) Dentist: I'm going to polish your teeth.
Patient: *Eh*?

Here, there is no suggestion that the particle functions to request confirmation in the same way as it does in (21a). The patient is simply requesting the dentist to repeat what she has said.

The fifth and sixth problems with the conceptualist account are also solved. The non-truth-conditional status of interjections, which is hard to explain on a fully conceptual account, is to be expected if they encode procedures which often fall on the non-truth-conditional side. And under a procedural account, there is no expectation that *ouch* and 'I feel pain' will be synonymous.

While solving these problems, the procedural account preserves the conceptualist intuition that there is a coded element to interjections, responsible for their language-specific nature, and Goffman's intuition that interjections are more than mere natural display. It also allows us to incorporate aspects of the functional treatment that Goffman proposes, by suggesting a plausible way in which the communicative content he describes might actually be communicated: via a combination of procedural encoding and inference.

However, one of the problems I raised with the conceptualist account remains. I claimed that it overlooks the fact that interjections seem to share with paralinguistic or non-linguistic behaviours the property of being partly natural (as well as partly coded). As yet, other than proposing that interjections might work by activating certain attitudinal descriptions, I have said nothing about this partly natural side, nor how it might be reconciled with the coded side. For while we intuitively regard words that encode procedural meaning (e.g. 'so', 'after all', 'however', 'moreover', 'I', 'he' etc.) as properly linguistic items, there remains a doubt as to the linguistic status of interjections.

Another issue that I have not yet addressed is the fact that interjections can constitute utterances in their own right in a unique non-elliptical manner; in such cases the higher-level explicature account proposed above would be problematic, since a higher-level explicature, by definition, takes an embedded proposition as its object.

In fact, the two issues are not unrelated, and a way of resolving both would be to see interjections themselves as working more in the manner of wholly natural phenomena, which can on occasion contribute to the construction of higher-level explicatures (when used by a hearer to flesh out a linguistically-

encoded logical form), but which contribute only to implicatures when used alone.

The issue of the linguistic status of interjections remains. It is, moreover, a question that has considerable bearing on the analysis being offered here. So far in the literature on the conceptual-procedural distinction, procedural meaning has only been attributed to linguistic expressions, and the question of whether a procedural account is compatible with non-linguistic status has not been addressed. I turn to this question in the next chapter. In the next section I look in more detail at the linguistic status of interjections.

6. Interjections and language: *are interjections part of language*?

In Chomskyan terms, knowing a language is having a mentally represented grammar, or I-language. However, we may also want to think of 'language' in wider terms. The human production and understanding of natural language is mediated by the grammar in conjunction with other cognitive systems. The ability to produce and understand language in this wider sense including the ability to perform various pragmatic processes of interpretation. It also includes the ability to attribute intentions and beliefs to others.

These observations are crucial in any attempt to answer question (3) above. For while interjections undoubtedly contribute to the interpretation of utterances, the same can be said for the whole range of natural phenomena discussed at various points in this thesis: although interjections may contribute to linguistic communication, it does not necessarily follow from this that they encode anything *linguistic*. For an interjection to be regarded as a part of language in the

narrow sense discussed above, the mentally-represented grammar must be involved: if interjections are part of language, they must encode linguistic information, i.e. that coding must be stipulated in the grammar.

Ameka summarises the conceptualist viewpoint on question (3) as follows:

> 'different interjections do have different degrees of integration within the linguistic systems of languages.[...] But the underlying commonality shared by all words which satisfy our characterisation of interjection is that they are linguistic signs.'
>
> (Ameka 1992, p. 113)

It is clear from the first part of this quote that although they see interjections as part of language, even the conceptualists allow for some borderline cases. Ameka argues there are three respects in which it might be argued that interjections are peripheral to language. These provide a convenient framework within which to approach question (3).

The first property of interjections that Ameka singles out is their 'paralinguistic' nature: 'there is no doubt that there is an intimate connection between interjections and gestures in general' (Ameka 1992, p. 112). Wierzbicka describes interjections as 'vocal gestures', which fits Goffman's intuitions that they are paralinguistic, and to a certain extent my own that they are partly natural as well as partly coded.

Wierzbicka does not, however, regard this as militating against a semantic analysis, and proposes to capture this by omitting the 'I say' component from her proposed conceptual structure (simplified as in Wierzbicka 1992, pp. 162-163):

(29a) *Ow*

I feel pain.

(29b) *I feel pain*

I say: I feel pain

I say this because I want to say how I feel.

This would remove interjections from the class of assertions, and leave them free to perform other speech acts—as expressives, for example. I find this an interesting proposal, and more in line with my own intuitions than other aspects of the conceptualist analysis. It seems to echo Kaplan's (and Searle's) descriptive/expressive distinction, in that (29b) describes (conceptualises) a feeling, while (29a) just expresses it.

Recall examples (23a) and (24a), repeated below:

(23a) *Aha*! You're here.

(23b) The speaker is surprised that I am here.

(24a) *Wow*! You're here.

(24b) The speaker is delighted that I am here.

A hearer of these utterances might well be led to construct higher-level explicatures such as (23b) and (24b) above. Given Wierzbicka's intuitions, and the framework discussed in Sections 4 and 5, the issue is whether he would also construct the higher-level explicatures in (30ab):

(30a) The speaker is saying that she is surprised that I am here.

(30b) The speaker is saying that she is delighted that I am here.

My intuition is that he would not, any more than he would construct (30a) and (30b) when a speaker says 'You're here!' and accompanies it with a surprised facial expression or a smile. This seems to support Wierzbicka's claim and might be taken as evidence that interjections are not part of language. However, Wierzbicka is not dissuaded from her conclusion: 'interjections—*like any other linguistic elements*—have their meaning, and … this meaning can be identified and captured in rigorous semantic formulae' (1992, p. 188—emphasis added, TW).

Wilkins disagrees with Wierzbicka's claim that use of interjections does not amount to 'saying'. On the contrary, he suggests, native speakers are happy to accept that some interjections are 'said', and presents evidence from his own informal survey to support this. He found that native speakers regarded (31a-b) as acceptable, but (31c-d) as unacceptable. These latter expressions are, he argues, better reported using the verb 'go' (31e-f):

(31a) 'Ouch!', she said.

(31b) 'Wow!', she said.

(31c) ??'Psst!', she said.

(31d) ??'Shh!', she said.

(31e) 'Psst', she went.

(31f) 'Shh!', she went.

He concludes that 'primary interjections are not merely vocal gestures' and 'interjections like *wow* and *ow* do have an 'I say' component in their decomposition, and may be regarded as illocutionary acts' (Wilkins 1992, pp. 147-8). He also claims his survey provides evidence that 'interjections that match the typical word phonology of English are regarded by native speakers as words' (Wilkins 1992, p. 148).

Here Wilkins touches on the second factor Ameka mentions: phonological atypicality. Wilkins' test in (31) suggests that there is a line beyond which items that are sometimes considered interjections (and are included in my original list) are not classified by native speakers as part of language. Vowel-less vocalisations such as *psst* and *shh* are two examples. Other examples from my introductory list include *brrr*, *hmm*, [|]—the dental click usually orthographically realised as *tut-tut* (or *tsk-tsk*), and *ahem*, often referred to as an interjection but in practical terms usually little more than an ostensive throat clear. *Oops* also fails to fit standard English phonotactics (English words do not begin with [ʊ]).[13] Similarly *ugh* differs from *yuk* in that the former ends in a velar fricative [x] that is not linguistically productive in English.

Essentially, Wilkins' argument is that since phonologically atypical interjections cannot be reported using the verb 'say', they are not part of the language. However, the situation is more complicated than he suggests, and the argument is not convincing. Not only can we use the reporting verb 'say' with many expressions which are clearly not words of the speaker's own grammar[14],

[13] I abstract away from a number of dialects in which the word 'up' begins with /ʊ/.

[14] There is evidence, in fact, that Wilkins is confusing the direct quotation use of 'say', with 'say' as in 'state' or 'assert' (i.e. the more technical term I think Wierzbicka has in mind).

in metalinguistic uses such as direct quotation, but 'go' is a perfectly acceptable verb with which to report linguistic utterances (32abc):

(32a) And so the kid would say, 'Blah blah blah?' [tentative voice with rising intonation] and his father would say 'Blah blah blah' [in a strong blustery voice], and they would go on like that.[15]

(32b) She looked at me and said 'moi, je déteste les dentistes'.

(32c) So he comes into the pub and he goes 'where's that money you owe me?'. 'What?', she goes, 'I don't owe you anything'.

Furthermore, a combination of the conceptual approach and Wilkins' claim that phonologically atypical interjections are not words would lead to considerable problems in accounting for the borderline expressions that Ameka alludes to. I don't think I am alone in having *yugh* [jəx] as well as *yuk* [jək] and *ugh* [əx] in my interjectional repertoire. Under Wilkins' account, *yuk* is part of language proper and communicates via its precise encoded conceptual structure: to suggest it does so solely because of this, however, leaves no account of *yugh*, which must surely communicate in a similar manner.

The third and final issue in deciding whether or not interjections are part of language is their syntax-independence and non-productivity. Interjections are, as it is often put, 'thrown' (interjected) into utterances. They exist on the edges of utterances, always separated off from the main clause and rarely integrated into intonational units. They do not inflect or combine with other morphemes to change word-class, and often stand alone as utterances in their own right,

[15] Clark and Gerrig (1990, p. 780), quoted in Wilson (2000).

seemingly without internal linguistic structure. If the crucial factor in deciding the linguistic status of interjections is whether or not the information they putatively encode is stipulated by the grammar, the fact that interjections operate independently of syntactic structure suggests they operate independently of the mental grammar.

In my introduction I stated that for the sake of argument I would assume that interjections represented a unified class. It should be clear by now, however, that this is not the case. As a further complication, consider (33ab):

(33a) At the Annual Dentist's Convention Mr. Pulley *wow*ed the audience with his encyclopaedic knowledge of gold teeth.

(33b) That is without doubt the *yuck*iest mouthwash I've ever tasted.

Wow and *yuk* are, of course, not secondary interjections: the linguistically productive expressions *to wow* and *yucky* (and *yummy*) are derived from the interjections rather than the other way round. This phenomenon complicates the picture even further, and the harder one looks, the more complicated it becomes.

Consider the utterances containing *eh* and *huh* in (21a) and (22a): although we cannot argue that these expressions are syntactically integrated, there is a sense in which they have to be 'thrown in' in a certain position in the overall utterance to perform the functions they do. With regard to phonology, recall Ameka's comment that interjections 'always constitute an intonation unit by themselves' (1992, p. 108). However, despite the comma in (34), *oh* could be the nucleus, or alternatively the pre-head of a larger intonational unit encompassing the whole phrase:

(34) Lily: That dentist's a complete sadist.

Jack: *Oh*, I don't know. (As in 'she isn't really'.)

They are such a disparate, non-unified group of expressions that the question whether 'interjections' are part of language may be impossible to answer satisfactorily; an adequate account of interjections should reflect this heterogeneity. It should also reflect the evidence I presented in this section, which suggests that many interjections are *not* part of language.

As mentioned above, the question remains whether, having argued against a conceptual and for a procedural approach, we can maintain the procedural approach in spite of this uncertain linguistic status. In the next section I focus on this *natural* side of interjections, and then suggest a way it might be reconciled with the *coded* side.

7. The naturalness of interjections

> "A Japanese picture of a hill both differs from and resembles a typical modern European painting of the same kind of hill. Both are suggested by and both 'imitate' the same natural feature. Neither the one nor the other is, in any intelligible sense, a direct outgrowth of the natural feature... The interjections of Japanese and English are, just so, suggested by a common natural prototype, the instinctive cries, and are thus unavoidably suggestive of each other."
>
> Edward Sapir (1970, p. 6)

Having looked at the coded element of interjections, it is now time to focus on their natural side. Recall Goffman's suggestion that interjections occupy a position on a continuum between display and language proper; interjections do indeed seem to lie half way between the two extremes. There is a sense in which they are partly natural, as well as partly coded.

In *The Expression of the Emotions in Man and Animals* (1872)[16] Darwin considers whether 'the sounds which are produced under various states of mind determine the shape of the mouth, or whether its shape is not determined by independent causes, and the sound thus modified' (*ibid.* p. 96). In describing the natural human expression of surprise he notes: 'Certainly a deep sound of a prolonged *Oh!* may be heard from a whole crowd of people immediately after witnessing an astonishing spectacle' (*ibid.* p. 97). He goes on: 'If, together with surprise, pain be felt, there is a tendency to contract all the muscles of the body, including those of the face, and the lips will then be drawn back; and this will perhaps account for the sound becoming higher and assuming the character of *Ah!* or *Ach!*' (*ibid.* p. 97). Despite the fact that interjections that express pain are language specific—English *ouch*, French *aïe*, Spanish *ay*, Finnish *auts*—they do all begin with the same mid-front vowel that Darwin describes as being naturally expressive of pain. Darwin's observations of how humans naturally express surprise and astonishment (and wonder) suggest that certainly *oh* arises out of a natural behaviour. And he notes other natural expressions of surprise: 'the dropping of the jaw and open mouth of a man stupefied by amazement' (*ibid.* p. 284); the fact that 'when thus affected, our mouths are generally opened, yet the lips are often a little protruded' (*ibid.* p. 285). Given these observations, *aha* and *wow* might also be viewed as developments out of natural behaviours.

When discussing the natural expression of disgust, Darwin says: 'With respect to the face, moderate disgust is exhibited in various ways ... by blowing out of the protruded lips; or by a sound as of clearing the throat. Such guttural

[16] All quotes are from the 1998 edition (edited by Paul Ekman—see bibliography).

sounds are written *ach* or *ugh...*' (*ibid.* p. 256). The interjection *yuk*, then, is closely related to the natural expression of disgust.

As Sapir points out in the epigraph to this section, this goes some way towards explaining why interjections, although not entirely involuntary reactions, *feel* so instinctive both to speaker and hearer. Standing alone in the kitchen, we do not utter 'I feel pain' if the kitchen knife slips, we utter *ouch*. In terms of interpretation, if you hear a spontaneous utterance of *ouch*, the evidence for that first layer of information, that the speaker is in pain, seems direct in a way that 'I feel pain' is not. In this respect, the continuum presented in Chapter Three reflects the intuitions behind Goffman's own proposed continuum.

We have already seen that in relevance theory, as well as degrees of explicitness at utterance level, there are degrees of explicitness at word level; a particular word may be used to express not exactly the concept it encodes, but another related concept which is more relevant in a given context. However, there is another way in which coding and inference may interact at the lexical level. While they are clearly properly linguistic, certain words appear to carry an extra element of 'showing', where the evidence provided for the first layer of information is more direct.

Onomatopoeic language is an obvious example (e.g. 'clink', 'clank', 'splash', 'sizzle'). In fact, iconic language generally is an example: stylised imitations of non-human sounds (e.g. *buzz*, *miaow*, *moo*, *oink*); also, stylised imitations of human sounds (e.g. *ha ha*, *tee hee*, *boo hoo*, *boo*, *hiccup*). In these last examples, there is an element of coding which separates them from clear instances of showing, such as laughing or crying, but also an element of showing, which separates them from clear instances of saying, such as 'I am amused' or 'I am

crying'. The link here between sound and meaning is not entirely non-natural or arbitrary. Indeed, we might argue that 'yuk', for example, originated as a stylised imitation of the natural expression of disgust discussed by Darwin above.

On this link between sound and meaning, Grice observed that:

> 'Any link will do...and the looser the links creatures are in a position to use, the greater the freedom they will have as communicators, since they will be less and less restricted by the need to rely on prior natural connections.'
>
> (1989, p. 296)

In stylised imitations of the type described above, and in onomatopoeic expressions generally, the link between sound and meaning is not *as* loose as in most other words since some element of the natural connection remains. The fact that some stylised imitations have been grammaticalised to the point where they are linguistically productive suggests the relation between coding and inference is even more complex. This point is illustrated in (35abc):

(35a) The bacon was *sizzling* in the pan.

(35b) The cows were *mooing*.

(35c) He *hiccuped* loudly.

This is not to suggest that there are degrees of coding, or to attempt to blur the distinction between coded and non-coded meaning. The suggestion is that there might be different *types* of coding. In the above examples, there is an iconic element, and the hearer is given more direct evidence of the first layer of

meaning than in examples of pure coding.[17] Since many interjections seem to be exaggerations or developments of natural expressions of emotion, they might also be regarded as stylised imitations, and iconic in some way; although for reasons discussed in the previous section, and in contrast with the examples in (35abc), they are not properly linguistic.

In fact, even some of those vocalisations which I have been treating as interjections, but which cannot be shown to be derived from natural expressions of emotion, are iconic to some extent. *Shh* does not convey emotion: but it could be argued that its voiceless quality, together with the fact that it can be uttered continuously, make it a particularly suitable sound—but not *word*—for urging someone to be quiet. The showing-meaning continuum, then, can be seen to apply at a *lexical* level as well, but in a manner that is somehow orthogonal to the lexical pragmatic processes discussed in the previous chapter.

More evidence that these expressions are located along a continuum is that there really does appear to be a gradual increase in stylisation/codification among them. This reflects the parallel drawn by Goffman (1981) between interjections and *ritualised* behaviours, in the ethological sense of that term. Consider *shh*, *shush* and 'hush'; consider the progression noted earlier from *ugh* to *yugh* to *yuk* to 'yucky'. Similar progressions can be seen from [ostensive throat clear] to *ahem* to the highly stylised [əˈhəˈhəm], or from [dental click dental click] to *tsk tsk* to *tut tut* to 'he *tutted* loudly'. Right down at 'word' level there appears to be a continuum from more direct to less direct evidence: from *showing* to *meaning*.

[17] I am grateful to the anonymous referee who pointed out some interesting data relevant to this topic in a comment on a previously published version of this chapter (Wharton 2003a). Cuxac (1999) describes the way in which deaf children raised by non-signing parents spontaneously develop 'iconic' signs, which are then used to communicate by both the children and the parents.

So to return to the three questions asked in my introduction to this chapter:

What do interjections communicate? I have argued that interjections communicate attitudinal information, relating to the emotional or mental state of the speaker. In some instances the attitude might be genuinely propositional: say, an attitude of questioning, or regret, or joy, or sadness, directed at an embedded propositional content. However, sometimes the mental state is directed not toward an embedded proposition, but toward a percept or object which is the cause of a feeling or sensation. In some instances, what is expressed is merely a feeling or sensation with no apparent cause. In all these cases, what is communicated may be extremely vague: in relevance-theoretic terms it will involve only a marginal increase in the *manifestness* of a very wide range of assumptions. Such vagueness is captured by the procedural account proposed earlier: the greater the range of attitudinal concepts activated by the procedure, the greater the vagueness.

How do interjections communicate? Interjections are partly natural and partly coded. As Goffman suggests, they fall at various points along a continuum between display and language proper; Goffman's own continuum, as I mentioned, appears to be rooted entirely in the notion of coding and increased codification. However, we can recast his intuitions in terms of the showing-meaning$_{NN}$ continuum introduced in the previous chapter. This captures the partly natural, partly coded nature of interjections, because whilst in one way interjections offer fairly direct evidence of the basic layer of information being communicated, in another their partly coded nature makes them somehow *less* direct than completely spontaneous, natural sounds. The continuum also allows to capture the heterogeneity and marginal linguistic status of the class in general.

Seeing interjections in this way, we should not be surprised that the attitudes they communicate are not always propositional. Nor should we be surprised that what they convey is sometimes too nebulous to be paraphrased in fixed conceptual terms: they are *partly* natural responses.

In fact, there is good reason to suppose that some interjections are derived from natural expressions of emotion, and for this reason the continuum may have diachronic implications. The element of stylisation or coding in interjections takes them beyond pure showing; this stylisation is also present in some aspects of language proper, at the meaning end of the continuum. Other interjections are not 'natural' in this sense, but may also be iconic—e.g. *shh*: these also fall somewhere between showing and meaning. With all interjections, the evidence provided for the first layer of information is more direct than with saying, but less direct than with entirely natural behaviours.

I have argued that the coded element of interjections is procedural, and that what is activated by the use of an interjection might be used by the hearer in a variety of ways. When combined with a sentence, it may function in a similar way to other natural phenomena, by encouraging the construction of higher-level explicatures.

In an utterance which consists of *just* an interjection, and expresses no explicit proposition, the obvious account would be to suggest that a hearer can only use the procedural information to derive implicatures: what the attitude is, what it is to, what the emotional/mental state of the speaker is. In this respect, interjections would pattern with paralinguistic and non-verbal behaviours generally; for while these might help a hearer construct higher-level explicatures when interpreting a linguistic utterance, they cannot contribute to explicit

communication when used alone as an ostensive stimulus. However, this account is slightly problematic. Since interjections have a coded element, it seems unsatisfactory to suggest that they only contribute to the implicit side of communication. I return to this issue in the next chapter.

Are interjections part of language? Since there is a continuum involving different combinations of natural and coded information, we would expect that it might be possible for expressions to move along it. In historical terms, when an interjection moves far enough along the continuum, it may become relatively productive ('to wow', 'yucky'), and some of its uses may be properly linguistic (verbs, adjectives etc.). When used as an interjection, though, given its similarities to paralinguistic phenomena, it seems to retain its independence from the mental grammar.

The answer, then, is no, interjections are not part of language; but the continuum does offer a framework within which they might be seen as existing on the edge of language, integrated to a greater or lesser extent: to use Goffman's expression—*semiwords*. This conclusion is further supported by aphasiological evidence of a dissociation between interjections and language proper. Goodglass (1993) demonstrates that interjections such as *ouch* remain within the repertoire of certain grave aphasics. If an individual can have interjections, but not have language, it is hard to see how the former can be viewed as part of the latter.

A question still remains, however: is there any further motivation for the claim that interjections, despite their non-linguistic status, might encode *procedural* information? Having argued that interjections are not linguistic, what light can be shed on interjections by an analysis so deeply rooted in *linguistic* distinctions? For evidence, I turn in the next chapter from the *semi*-natural to

uncontroversially natural phenomena such as facial expressions and spontaneous expressions of emotion. I argue that a sub-set of these behaviours have an inherent signalling function: they are, in effect, *natural codes*.

This has two implications: firstly, it suggests that Grice's natural/non-natural distinction, presented in Chapter One, is not exhaustive; secondly, it suggests that we need an account of precisely what kind of information these natural codes encode. If this turns out to be non-translational too, it should provide us with some motivation for pursuing the procedural account of interjections offered above and, moreover, may take us a step further toward understanding the processes that underlie the interpretation of complex ostensive stimuli.

Chapter Five

1. Natural codes

'There is no evidence about precisely what type of information is conveyed when, during an on-going social interaction, one person sees a facial expression of emotion on another's face.'

(Paul Ekman 1989, p. 159))

1.0 Codes, honeybee-dances and smiles

One of Grice's most lasting achievements was to provide an alternative to the code model view of communication. Under this latter approach, an utterance is a signal which *encodes* the thought or message a communicator wishes to communicate: in order to retrieve the speaker's 'meaning', all the hearer need do is *decode* the signal the speaker has provided back into the original thought or message. Construed in this way, linguistic communication (in fact, communication of any sort) works according to broadly the same principles as semaphore, or Morse code.

The assumption that human communication was a matter of coding and decoding was one of the key ideas underlying the *semiotic* programme (Peirce 1897, 1903, de Saussure 1916/1974, Vygotsky 1962). Indeed, this programme proposed that most aspects of human life—language, customs, rites etc.—were best analysed as systems of signs, or fundamental codes, which underpin and facilitate every type of human social and cultural interaction.[1] As we have seen (and as we shall see), many approaches to 'meaning' and communication come heavily laden with semiotic baggage.

[1] See Sperber 1996 for an alternative, naturalistic approach to culture, within which the socio-cultural domain is partly analysed in terms of an accumulation of individual psychological phenomena.

Broadly speaking, there appear to be two different answers to the question of what actually constitutes a code. On the one side, there is the strict semiotic sense: a code in these terms is a system which pairs a signal with a message, enabling two information-processing systems to communicate. 'Zoosemiotician' Thomas Sebeok, aspects of whose account of human communication I discuss in Section 2 below, summarises Shannon and Weaver's classic (1949) model as follows:

> 'One system, a source, influences another system, a destination, by dispatching alternative signals that are carried in the channel connecting them. The information source is conceived as producing one or more messages which must be transformed, or encoded, by a transmitter into signals which the channel has the capacity to carry; these signals must finally be transformed, or decoded, by a receiver back into messages which can be accepted by the destination.'
>
> (1972, pp.12-13)

On the other side, there is the notion of code most often used in the social sciences, and at least as common in ordinary linguistic usage. This is the notion of a code as a collection of rules, regulations or *conventions*: self-perpetuating regularities in the sense of Lewis (1969). Thus, we might speak of a code of law or a code of politeness, the Christian code or a code of ethics.

While the two notions are clearly distinct, they do not have to be seen as mutually exclusive. Both, for example, are recruited in explanations of human language. In the first sense, language might be seen as a system that pairs signals (sentences) with messages (meanings). In the second, many people (including Lewis, and Grice himself) have sought to analyse language as a set of signalling

conventions. Grice, for example, saw aspects of his natural/non-natural distinction reflected in the distinction between natural and *conventional* signs.

The sense in which I use the word 'code' in this chapter is the first one: the strict semiotic sense. Furthermore, I use it to the exclusion of the second sense. Given the Chomskyan perspective on language adopted in this thesis, it seems inappropriate to talk about language as a set of socially agreed-upon conventions in the same breath as an innate language faculty, or a *Universal Grammar*, which constrains the form of possible human languages. William Lycan sees major problems inherent in Grice's and Lewis' attempts to characterise literal meaning as a convention to use certain expressions with certain intentions:

> '...most sentences of a language are never tokened at all; since hearers instantly understand novel sentences, this cannot be in virtue of pre-established conventions or expectations directed on those sentences individually.'
>
> (1991, p. 84)

While many will disagree with any attempt to sever the link between human language and convention, few would advocate talk of *conventions* when discussing non-human animal communication systems. However, in the first, strict semiotic sense, codes they most certainly are. In that regard, the existence of what I will call *natural codes* is uncontroversial.

Consider honeybees: the honeybee performs a complex dance in order to indicate to its conspecifics information relating to the location of a source of nectar (von Frisch 1967). The dance can 'transcend the here and now and ... make reference to distant temporal and spatial variables in the environment rather than only to the immediate surroundings of the signaller' (Allen and

Bekoff 1997, p. 108). The distance of the food source from the hive is indicated by the length of the dance; direction away from the hive is conveyed by the orientation of the 'waggle' component of the bee's dance in relation to the position of the sun. Recent research (Dreller and Kirchner 1993, 1994) has suggested that there may well be an auditory, as well as visual, dimension to the bee's dance.

Would we want to characterise the 'meaning' the dances have as $meaning_N$ or $meaning_{NN}$? Recall the tests from Chapter One: Firstly, is the meaning carried by these dances factive or non-factive? It seems fair to suggest that it is factive: the fact that the honeybee has performed the dance means that the nectar is there;[2] secondly, is there any evidence to suggest that the interpretation of the dance relies on the deployment and attribution of intentions? As far as we know, there is none. The dances of honeybees, then, mean naturally.[3]

---

[2] The factivity test breaks down somewhat in the case of bee-dancing. If the bee makes a mistake—as I'm sure bees occasionally do—then there is at least a sense in which the dance *still* meant 'nectar-at-location *x*'. The law-like link between, for example, black clouds and rain (in which if it doesn't rain then those black clouds can't be said to have meant rain anymore) appears not to hold in the case of bee-dancing. Furthermore, I am claiming here that the factivity test might be used to sharpen our intuitions on the *nature of the meaning carried by the dances themselves*, and I'm aware that this does not sit entirely comfortably with my earlier remarks about how Grice's examples might best be seen as utterances *about* black clouds or spots (or bee-dances)—see pp. 27-29. I would add, however, that if we try to fit bee-dancing into Grice's natural/non-natural dichotomy, I don't see that we have any choice but to view it as an example of *natural* meaning since, as I go on to say, it seems fairly clear that bees do not possess the kind of higher-order intentional ability to *mean non*-naturally (in the sense described by Grice). Such abilities are a pre-requisite for the existence of non-natural meaning.

[3] I had originally planned to include another example from the world of non-human animal communication at this point: the alarm calls of vervet monkeys, for instance. However, the fact that vervets do not automatically emit an alarm call on seeing a predator (Cheney and Seyfarth 1990) might be seen to complicate the issue somewhat. It has been claimed (by Dennett 1987) that this is evidence that their calls are *intended* for other vervets. Notice, however, that the vervets' calls would still be factive if they were *only* emitted in the presence of a predator; what

There is, nonetheless, a sense in which it is plainly unsatisfactory to see the meaning of the honeybees' dances as entirely parallel to paradigmatic examples of Gricean meaning$_N$ such as 'those black clouds mean rain'. There is nothing *coincidental* about the fact that a honeybee's dance 'means' something to another honeybee: the function of the honeybees' dance is to indicate. It is not the function of black clouds to convey the information that it is going to rain.

In fact, this observation is neatly reflected if we apply some of Grice's tests for meaning$_{NN}$ to honeybee-dances. The results, which in the examples in Chapter One illustrate so neatly the 'reasonably clear intuitive distinction' he sought to demonstrate, become strangely unreliable in (1ab):

(1a) What is meant by the honeybee's dance is that there is nectar at location$_X$.

(1b) That honeybee dance means "nectar-at-location$_X$".

At least in the case of honeybee dancing, it appears that forms of paraphrasing previously only appropriate to describe cases of *non*-natural meaning, are also appropriate in cases that other tests suggest are cases of *natural* meaning. Why the tests become unreliable is unclear, but it seems to me that the most plausible explanation is that our intuitions concerning (1ab) simply reflect the fact that the dances of honeybees are inherently communicative: they are *coded* signals.

---

would be required to demonstrate that they were not factive would be cases in which vervets call in the *absence* of the appropriate predator: i.e. in cases of deception. Furthermore, even if there *were* evidence of deception among vervets (and, anecdotally, I have heard that there is) this would still not count as evidence that intentions play a role in either their deployment or comprehension. There are a variety of creatures that either feign injury (plovers) or play dead (snakes) when confronted by a predator, and it remains unclear whether this amounts to deception of the 'intentional' variety (see Hauser 1996, pp. 586-594 for discussion).

Honeybees don't 'mean' as in 'intend', but something is surely meant (in one sense of the word) by their dances (hence von Frisch's pioneering work). In this sense, then, Grice's dichotomy is not exhaustive.

It might be objected at this point that Grice's 1957 paper was not remotely concerned with non-human animal communication, and that the above observations are so utterly unrelated to the distinctions he was discussing as to be irrelevant. However, I do think that the complications raised by (1a) and (1b) have implications beyond the rather trivial observation that Grice's natural/non-natural dichotomy fails to accommodate the dances of honeybees. My reason is this: *the very same complications arise if we apply the tests to certain human behaviours*. In particular, they arise with a subset of human behaviours that are, to all intents and purposes, natural signs: facial expressions such as smiles, for example.

Consider again the three natural behaviours discussed in Section 3 of Chapter One: (involuntary) smiles, crying and shivering. These are natural signs that a person is happy (or at least not a threat), distressed and unhappy, or cold. Recall the tests again: the 'meaning' being carried in all three cases, for example, is factive: a spontaneous smile means that person *is* happy—*now*; crying means that a person *is* distressed or unhappy; a spontaneous shiver means that person *is* cold. This is borne out in (2abc) below:

(2a) The fact that he is smiling means he is happy.

(2b) The fact that he is crying means he is unhappy or distressed.

(2c) The fact that he is shivering means he is cold.

As we have already seen, the question of whether or not the attribution of intentions might be said to play a role in the interpretation of human natural behaviours is complicated by the fact that in a species that is aware[4] of their (involuntary) production, they might be deliberately shown to make evident an informative intention; furthermore, such behaviours might be exaggerated, developed or faked in communicative situations.

However, there is clearly a sense in which the message carried by these natural behaviours can be said to be derivable without reference to the intentions of the person responsible for the behaviour: that is *why* we regard them as natural. It doesn't contradict my earlier position—that the deliberate showing of natural behaviours might be considered intentional acts—to acknowledge the fact that natural behaviours convey information *whether or not* they are deliberately shown.

Despite the apparent 'naturalness' of all three behaviours, those Gricean tests that yield interesting results when applied to honeybee-dances also yield interesting results when applied to smiling, which behaves differently under the tests to shivering and crying. There seems to me to be a sense in which (3a-b) are acceptable in a way that (3c-d) possibly, and (3e-f) certainly, are not.

[4] I don't want to become embroiled in a debate about whether or not bees are 'aware' or even 'conscious' (whatever that means). Suffice to say, it is doubtful that they have the cognitive abilities to reflect on the content of their coded signals to the same degree as humans (though see *fn.* 29 in Chapter Six).

(3a) His smile means "I am happy"

(3b) What was meant by that smile was that he is happy.

(3c) ?His tears mean "I am distressed or unhappy'.

(3d) ?What was meant by his tears was that he is distressed or unhappy.

(3e) *His shiver means "I am cold".

(3f) *What was meant by that shiver was that he is cold.

And remember again that I am talking about *spontaneous* smiles here. It could be argued that since a fake smile can be used to mean non-naturally, the intuitions that the tests rely on are confused. I do not think this is the case. Consider a situation in which someone uses a fake, forced smile to $mean_{NN}$ something like 'I am not amused' (that slightly sardonic smile, which is often accompanied by a monotone *ha-ha*). In this case (4ab) *would* be appropriate paraphrases:

(4a) Her (sardonic) smile means "I am not amused".

(4b) What was meant by her (sardonic) smile was that she was not amused.

And it seems clear to me that (3ab) are acceptable in a manner distinct from this, a manner more closely akin to the cases in (1ab).

Just as the dances of honeybees 'mean' in a stronger (or at least different) sense than black clouds, so a spontaneous smile 'means' in a stronger (or different) sense than crying or shivering. And as with honeybee dancing, the reason for this, I claim, lies in the adaptive functions of the behaviours themselves.

1.1 Signs and signals

In order to clarify which instances of information transmission in the non-human animal world are to be regarded as communication, and which are not, Marc Hauser (1996, pp. 9-10) distinguishes between two ethological notions: *signals* and *signs*.[5] *Signals* are those behaviours that have been 'designed', in the sense that they are adaptive, to convey information. The dances of honeybees and, say, the alarm calls of vervet monkeys are signals: their primary function is to convey information.

Whilst they may be highly informative, *signs*, on the other hand, do not have a signalling function. Hauser provides two examples: In the first, he conjectures that forest monkeys might use the presence of chimpanzee nests to avoid chimpanzees, and hence predation. However, the evolutionary function of chimpanzee nests is not to inform forest monkeys of the presence of predators. In the second, as a result of regular travels across dusty soils, predatory species such as lions and pythons might leave traces of their presence. Certain prey species might learn that particular traces are associated with danger whereas other traces are not. The traces, however, cannot be said to have a signalling function. This, then, is an ethological version of the distinction I alluded to in the introduction, between indicators the function of which is to indicate, and those which have no such function.[6]

---

[5] Hauser draws a further distinction between 'signs' and '*cues*'. The latter are communicative phenomena such as sexual ornaments and warning colours, which are permanently 'on'. The distinction has no bearing on the discussion in hand, but it is worth noting that these too are natural indicators.

[6] Of course, this kind of exploitation of natural signs may well turn out to be highly adaptive. Indeed, it may be selected for and lead to a kind of evolved cognitive reflex in the form of a correlation between a particular feature of the environment and a certain reaction in a given organism. I'm thinking of the 'fear-reflex' most people feel when confronted by, for example, a

It is tempting to view the ethological notion of a sign and Gricean natural meaning as entirely parallel: the nests indicate the presence of chimpanzees whether or not the forest monkeys take them that way; the tracks of lions and pythons indicate danger to certain prey species (if they are noticed). We can equally imagine these signs being interpreted by a human naturalist in the field (or perhaps better, a native tracker) along the lines of *those black clouds mean rain*. However, as we have seen, it would be a mistake to draw such a parallel. The Gricean natural/non-natural distinction does not fully accommodate the distinction between signs and signals because some signals are natural.

However, the ethological signal-sign distinction does capture the distinction alluded to above between smiles and shivers. Smiling, after all, evolved as a signalling activity (van Hooff 1972, Fridlund 1994, Ekman 1999): its function is to indicate, or carry 'meaning'. The function of the shiver response, on the other hand, is to generate heat by rapid muscle movement. In ethological terms, smiles are signals, and shivers are signs:[7, 8] natural behaviours, then, do not all work in the same way.

---

large snake. Though I'm pre-empting the discussion below to a certain extent, we would not want to call the relationship between the sign (the snake) and the reflex (fear) a 'code', for it is still not the function of the *snake* to indicate that there is a snake in the vicinity; rather, it is the function of the *reflex* to keep the person safe. As ever, the issue is much more complex than I am presenting, since the reason you notice the snake in the first place may be because it has conspicuous markings, the function of which—since the snake is venomous—is to give the snake away.

[7] The reader will have noticed that I am not discussing the function of crying. There are two reasons for this: firstly, I'm not as convinced by the results of the tests in (3a-f) as applied to crying as I am by the results as applied to shivering and smiling; secondly, the dissociation between crying and tears makes it hard to say with any authority exactly what the adaptive value of crying is (as distinct from shedding tears). See Darwin 1872/1998, pp. 164-175, also Hauser 1996, p. 469 for discussion.

In the ethological literature, non-human animal communication systems are often referred to as codes (see Bradbury and Vehrencamp 1998, pp. 456-457 for one example). Parallel to this, the evolutionary link between *signal* and *message* in behaviours such as smiles suggests they too are best analysed as coded behaviours: *natural codes*.[9] One difference, then, between the interpretation of smiles and shivers would be that some of the cognitive processes responsible for the interpretation of smiles would not be the same as the inferential processes responsible for the interpretation of shivers, and other *signs*. The interpretation of smiles and other spontaneous expressions of emotion will (at least initially) be automatic and sub-conscious, more typical of the immediate coding-decoding responses so typical of non-human animal communication.

---

[8] Herb Clark (1994) makes a similar distinction, between the 'meaning' conveyed by natural signs—which he calls *symptoms*—and 'the meaning of certain deliberate human acts... *signals*'. There are interesting parallels between some of the issues discussed in this paper and those covered in Clark's book. However, the 'sign-signal' distinction introduced above and Clark's 'symptom-signal' distinction are not co-extensive. Firstly, the sign-signal distinction *I* make is a distinction *within* the category of phenomena that Clark calls symptoms (see *fig.* 3 on page 192—the top node might be labelled *symptoms*). Secondly, as I have tried to show in Section 2, many spontaneous, involuntary 'symptoms' that do not have an indicating function—that are *signs* in my terms—can still be deliberately *shown* in an act of intentional communication. Although we might describe these as 'deliberate human acts', I don't think Clark would want to call them signals (in his terms).

[9] Von Frisch (1967) proposes that the honeybee's dance has its evolutionary origins in primitive bees' pre-flight *intention* movements ('intention' in the ethological sense of reliably correlated with a certain course of action or behaviour—see *fn.* 2, p. 9 above) which became refined, stylised and stereotyped over time. Signals, then, may well have their evolutionary origins in signs and this—as far as I understand it—is a fairly standard ethological account of one of the ways in which coded behaviours (signs) might evolve from what was (to all intents and purposes) Gricean natural meaning. In the final chapter, I present an account of how 'intentions' (in the 'rich, philosophical sense'—see *fn.* 2, p. 9 again) might have played a role in the evolution from meaning_{N} to meaning_{NN} (as distinct from purely coded 'meaning').

In fact, although Shannon and Weaver's model was developed with engineering problems in mind, it seems a highly appropriate blueprint on which to model animal communication systems. The stimulus provided by the transmitting animal is the signal that encodes the message. The cognitive or affective state activated in the receiving animal is the decoded message. Of course, the encoding and decoding of the honeybees' dance—as in non-human animal communication generally—is an automatic process. It occurs without either the sending or receiving animal consciously recognising that the signal *means* anything.

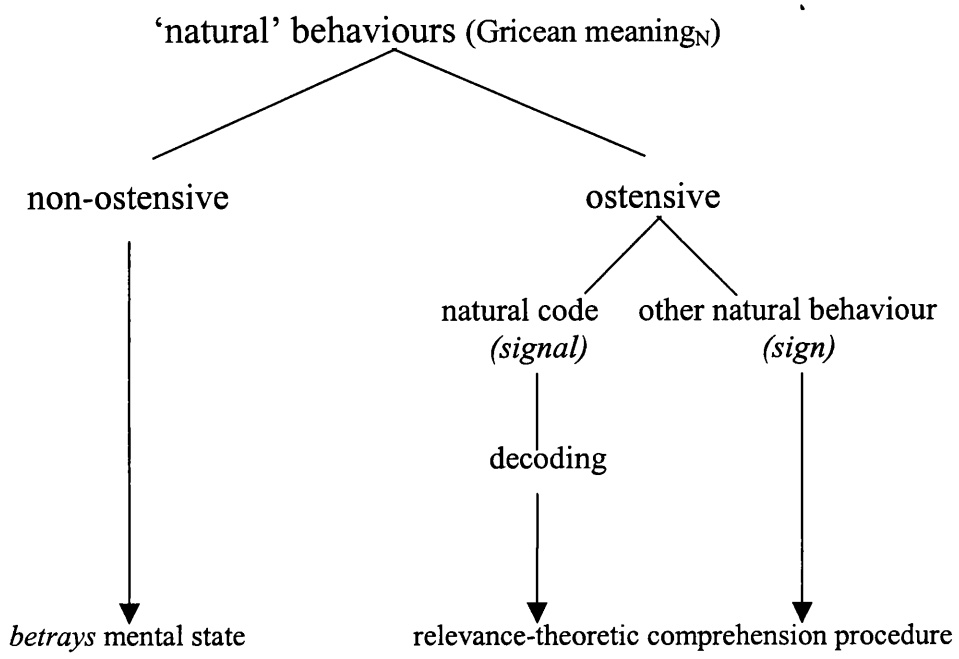
In one sense, human natural codes are the same. We read facial expressions, for example, automatically. They activate in us a particular mental or emotional state that correlates with the mental or emotional state—surprise, delight, anger, fear—of the communicator. In another sense, however, they are not. The automatic reading of facial expressions needs to be reconciled with the fact that humans *can* reflect on the content of signals, and, what is more, know that others can reflect on them too.[10] For when natural coded behaviours are put to use in ostensive-inferential communication, the automatic decoding processes that govern their interpretation are supplemented by other equally specialised automatic—but this time inferential—processes that govern the search for relevance. If we integrate the account given in Section 3 of Chapter One with the observations made here, we can represent the new picture as in *fig. 3* below: [11]

---

[10] Sperber 2001 notes that the fact that humans have the ability to reflect on the content of signals leads to differences in not only *how* humans communicate, but *why*. (I return to this in my final chapter.)

[11] To complete the diagram there should perhaps be a version of the signal/sign distinction under the 'non-ostensive' node. I am, however, reluctant to include one. There are two reasons for this. Firstly, as I pointed out in my introduction, I am taking the domain of pragmatic

fig. 3



The work of Paul Ekman (1989, 1992, 1994, 1999) suggests that there is a whole range of spontaneous facial expressions that have evolved in humans to reflect a signaller's internal state, and might thus be analysed as natural codes; 'these expressions have been selected and refined over the course of evolution for their role in social communication' (Ekman 1999, p. 51). Ekman's claims that these expressions reflect the existence of underlying basic, universal human emotions, which are to some degree at least biologically inherited, have been criticised (see Fridlund 1994, Russell 1994). Fridlund's *behavioural ecological*

---

principles to be intentional—ostensive—communicative acts. The focus of this chapter, therefore, is everything under the 'ostensive' node, and it is there that I am keen to flesh out the details. Secondly, and less obviously, though I would happily concede that 'signs' (such as shivers) do belong under the non-ostensive node, I am not sure that 'signals' (such as smiles) would sit there very comfortably. If I happen to see someone smiling to themselves, and interpret their putatively non-ostensive behaviour as showing that they are happy, there is still the question of why they are smiling in the first place. A (somewhat sceptical) colleague once remarked to me that in the same way that we might offer 'people shiver when they are cold' as an explanation of why people shiver, it was enough to offer 'people smile when they are happy' as an explanation of why people smile. I hope I have explained why it is not.

view, for example, stresses the manipulative communicative function of facial expression, whilst denying they are reliably correlated with underlying 'emotion'. Among the evidence he presents in support of this position are data from experiments on 'audience effects' in human smiling (Kraut and Johnson 1979). In these experiments researchers monitored the smiles of people involved in various activities—ten-pin bowling, spectating at an ice-hockey match—and found that, in general, people smile more for the benefit of others than themselves. During a ten-pin bowling match, for example, 'subjects rarely smiled while facing the pins, but did so frequently when they pivoted to face their friends in the waiting pit' (Fridlund 1994, p. 153).[12]

Hauser (1996, pp. 495-6) does not regard the two approaches as mutually exclusive: '...the debate actually confuses two levels of analysis. Whereas Ekman's work has generally focussed on the *mechanisms* underlying facial expression (e.g. changes in physiology, brain state), Fridlund has considered the *function* of facial expression'. One thing is clear, however; despite the disagreement over the putative role of actual emotion in the production of facial expressions, smiles and other spontaneous expressions of emotion differ crucially from facial reflexes such as eye-blinks or sneezes, or other non-communicative behaviours such as shivers: their function is to signal.

[12] Robert Provine, in his recent book *Laughter: A Scientific Investigation* remarks that there is 'a strong association between smiling and social motivation and an erratic association with emotional experience' (2000, p. 45).

2. What do natural codes encode?

'A six-word dictionary for grasshoppers (*Acrididae*):
Signal I: It is fine, life is good;
Signal II: I would like to make love;
Signal III: You are trespassing on my territory;
Signal IV: She's mine (of the female of course);
Signal V: Oh, how nice it would be to make love!
Signal VI: How nice to have made love!'
(adapted from Moles 1963, pp. 125-126)

2.0 Sebeok's *analogue* codes and Peircean *indices*

Thomas Sebeok (1972) investigates the different types of coding humans might use in linguistic communication, and proposes that there are two kinds: *digital* and *analogue*. His distinction was intended to reflect what he termed 'rational' and 'emotional' (1972, p. 10) human communicative content, and since much of the work on human facial expression in particular appears to be closely linked to 'emotional' or 'affective' content, it is worth investigating how much Sebeok's distinction can tell us about human natural codes.

The analogue-digital distinction exists in a variety of guises.[13] Essentially, it is the difference between codes or systems within which the repertoire of signals used to convey a message is either—in the case of analogue codes—graded, blended or continuous, or—in the case of digital codes—discrete or discontinuous. Within a graded system the boundaries between the signals cannot be demarcated, whereas within a digital one they can.

Recall the analogy from Chapter Three. An analogue system works as follows. Imagine a variable of some physical quantity: the pressure of a certain gas in a certain system, for example. This variable is related to another variable,

[13] Hauser (1996, p. 54) credits the cognitive ethologist Peter Marler with being the first to propose the distinction as applied to communicative systems.

say the needle in a pressure gauge, in such a way that the variations in the former are in a proportional relationship to the latter.[14] As the pressure in the system rises, so the needle on the pressure gauge rises; as the pressure falls, the needle falls. The movement of the needle is *analogous* to the rising and falling of the pressure, and the continuous fluctuation of pressure is reflected in the continuous movement of the needle.

In a digital system, the continuous flow of data—in our example the continuous fluctuation of pressure—is represented in terms of discontinuous or discrete units. In order to achieve this, the rise and fall of pressure is reanalysed by some converter-mechanism into these units: in short, the data are not *measured*, but *counted*. Rather than assessing the pressure by means of checking the quivering needle, the engineer consults a numerical, i.e. digital read-out.

The principal advantage of digital over analogue systems is one of accuracy: digitally encoded information is 'all-or-none'; analogue information is 'more-or-less'. Before recent technological advances in the recording industry, studios employed analogue recording systems using magnetic tape. The sound-waves produced by voices or instruments were converted by microphones into analogous waves of electrical voltage, and stored on magnetic tape as further sets of analogous waveforms, this time magnetic.[15] In modern computerised recording, the sound-waves are still converted by a microphone into voltage, but the electrical waves are re-analysed by a converter into discrete, digital (in this case binary) units. Again, accuracy is the principal advantage: the sound is a cleaner, supposedly more faithful representation of the original. Furthermore, the

[14] The pressure gauge analogy is a development of one presented in Bolinger (1983).

[15] The technical term is *magnetic flux density*.

digital representation allows much wider scope for modifying the original signal, i.e. altering the tuning or timing of individual notes (or part of notes).

However, just as words have their limitations, so do digital recordings. Modern—entirely digital—computerised recording studios employ state-of-the-art software to recreate the characteristic inaccuracies of analogue recording systems. Principally, this involves the reintroduction of extraneous 'noise' or 'hiss' to the sometimes antiseptically clean digital recording. They are thus (hopefully) imbued with some of the warmth and immediacy of older analogue recordings.[16]

One property taken to be one of the defining characteristics of human language is that it is a digital, combinatorial system. The discrete units—words—can be combined into larger structures which have properties that are distinct from the properties of the elements, and a product of the way in which the elements are combined according to recursive rules. Language is therefore 'infinite' as well as 'discrete',[17] and humans are capable of producing (and understanding) an unlimited number of distinct combinations, expressing an infinite range of properties. Within an analogue system there are no discrete elements to rearrange: the only way to distinguish a wide range of 'meanings' is to discriminate ever tinier and more subtle differences in the continuous signal.

As we have seen in examples throughout this thesis, the potential of analogue and digital codes to complement each other is nowhere more clearly illustrated than by human linguistic communication. Consider as further examples (5ab) and (6ab) below:

[16] This is one, but only one, of the reasons the old songs always sound best.

[17] To paraphrase Chomsky's (1988, p. 169) famous terminology.

(5a) Lily: Has John arrived?

(5b) Jack: John has arrived.

(6a) Lily: Has John arrived?

(6b) Jack: (smiling happily, in a pleased tone of voice) John has arrived.

By fronting the auxiliary in (5a), Lily forms an interrogative. One aspect of the difference in intended meaning between her utterance in (5a) and Jack's in (5b) is indicated by placing the discrete units of language in a different order. But consider Jack's replies in (5b) and (6b). In terms of linguistically encoded content, these utterances are identical. However, Lily would certainly be led to interpret them differently. Crucially, this difference in meaning is not achieved by digital means—Jack's smile and affective tone of voice in (6b) are in some sort of proportional or analogous relationship to the amount of affect he intends to convey: Lily reads his emotional state more in the manner of the engineer consulting the analogue pressure gauge, than the digital read-out. Depending on the breadth of his smile and the tone he uses, she might decide he is mildly pleased, quite happy or absolutely thrilled. Furthermore, the extent to which Lily can interpret these degrees of happiness depends not on her knowledge of any digital code, but on her ability to discriminate subtle (sometimes tiny) variations in his tone of voice, much as the engineer studies the quivering needle; we have already seen how such behaviours help speakers and hearers calibrate the appropriate sense of a given concept and feed into the relevance-theoretic comprehension procedure.

Of course, Jack might have chosen to try and convey this information digitally, and uttered 'I am mildly pleased that John has arrived' or 'I am quite

happy that John has arrived' or 'I am absolutely thrilled that John has arrived'. Notice, however, that utterances such as these produced in an entirely neutral tone of voice would sound extremely strange, and it is unlikely Jack would be taken to have communicated his feelings very effectively.

As well as discreteness, among the other distinguishing features of human language listed by Hockett (1959) was *arbitrariness*—de Saussure's (1916/1974) 'l'arbitraire du signe'. In fact, this notion lines up with that made by the father of modern semiotics Charles S. Peirce's notion of a *symbol*, which he distinguished from *icons* and *indices*. An iconic representation is one in which the relationship between the object and the representation is one of resemblance: a picture of a dog running is iconic of a state of affairs in which a dog is running. An icon 'has no dynamical connection with the object it represents; it simply happens that its qualities resemble those of the object' (2, p. 299).[18] An indexical representation is one in which the relationship is physically (or temporally) proportional or causal—the analogue pressure gauge described in the previous section is a good example: 'it is physically connected with its object...they make an organic pair' (2, p. 299). A symbolic representation is one in which the relationship between the signifier and the signified is governed by some social convention, tacit agreement or 'conventionalised'—the word 'dog' meaning dog: in other words, arbitrary. 'The symbol is connected with its object by virtue of the ... symbol-using mind, without which no such connection would exist' (2, p. 299).

Sebeok attempts to draw parallels between Peirce's distinctions and the analogue-digital distinction. His idea is that 'the most interesting thing about the property of arbitrariness is this: that it is a logical consequence of digital

[18] Quotes from Feibelman's '*An Introduction to Peirce's Philosophy*' (pp. 91-92); references to Peirce's '*Collected Writings' Volume 1-6.*

structuring in the code' (1972, p. 25); in effect, that it doesn't matter what the discrete units into which the continuous data are reanalysed are, nor what they are called: the discreteness alone suffices to make the code arbitrary. To the extent that Sebeok is right about digital lining up with arbitrary,[19] we might look at the extent to which analogue encoding lines up with the Peircean notions of icons and indices, and explore the extent to which (if at all) it helps us in our characterisation of natural codes.

It seems to me that analogue coding lines up most closely with notion of an index. Not only does it seem intuitively clear that a *picture* of a dog running is not a coded representation of a dog, but simply a likeness, but the causal and temporal links of Peircean indices reflect the kind of proportional relationship mentioned earlier as characteristic of analogue coding. This observation is reflected in a number of previous approaches. Jakobson and Halle (1956, p. 11) describe the expressive features of speech as 'physiognomic indices'. In an analysis with strong echoes of Peircean indices, Bolinger (1983) describes intonation generally as exhibiting properties of 'dynamic indicators' (though it is entitled 'The inherent *iconicism* of intonation'—see bibliography).[20]

The concepts of analogue codes and indices are certainly useful notions. The majority of human natural codes do indeed appear to work along analogue lines, and the notion does at least take us some way towards understanding how they

---

[19] Are there, for example, arbitrary *analogue* codes?

[20] An anonymous reviewer of the paper from which this chapter is largely an adaptation (Wharton 2001) objected to my close correlation of analogue encoding with indexicality. S/he argued that there need be no proportional relationship between the sign and the object in cases of indexicality, whereas in cases of analogue encoding there must. It may well be that the two notions do not match up exactly, but I hope that my comparison is at least germane, given that the causal and temporal links between sign and object in cases of indexicality do at least reflect those at the heart of analogue codes.

are interpreted. However, if we are to answer the question of what kind of information human natural codes encode, we must go further and say what analogicity and indexicality mean in cognitive terms. There are also, I suggest, problems inherent in any account that attempts to analyse natural codes in general solely in terms of analogue codes and indices.

Firstly, semiotic accounts rely entirely on a code model of communication. A coding-decoding model, however, is as *in*appropriate a framework within which to characterise the intention-driven communicative interaction of humans, as it is appropriate to characterise the dances of honeybees. Furthermore, semiotic approaches aspire to providing a 'general' theory of communication (albeit in terms of coding and decoding). As we have seen, there are two distinct types of communication: coded and inferential. Whilst it is true that in human communication these two types interact, neither should be regarded as more general than the other, for 'communication' does not depend on either: pure coding-decoding requires no inference, and inferential communication can take place in the absence of any pre-existing code.

Secondly, the idea that natural codes must always be analogue is too restrictive. While the honeybee's dance functions along largely analogue lines, codes of the kind used by vervet monkeys (Cheney and Seyfarth 1990) and certain birds (Okanoya 2000)—clearly natural codes in terms of the distinctions presented here—exploit discrete, digital signals. Thirdly, the idea that natural codes are indices is not restrictive enough: although in a sense all 'natural' indicators are indices (the etymological similarity is not accidental), it is not the case that all indices are codes. Recall the example of the mechanical system. The amount a particular pipe *bulges* is also in an indexical relationship to the rising of

the pressure; the engineer might indeed choose to assess the pressure system by checking the bulging seams. However, it is not the function of the bulging pipe to convey the information it does. It is a sign, not a signal.

The semiotic notions of analogue codes and indices are useful tools. They describe properties that human natural codes appear to have, properties which should be reflected in a satisfactory analysis. However, something more is required to say what the information encoded in a human natural code looks like in cognitive terms.

2.1 The conceptualist approach to facial expression

Wierzbicka (2000) also discusses the 'natural' properties of facial expression, and comes to similar conclusions regarding their inherent indexicality: 'those (if any) which are universally interpretable may have a "natural", i.e. iconic or indexical basis' (2000, p. 178); 'the basis for decoding lies either in similarity...or in co-occurrence' (2000, p. 156). Another area of common ground between her analysis and the one being developed here is that the 'naturalness' of certain human facial expressions does not preclude their being coded signals. As we have seen, however, I am reluctant to call *iconic* representations coded signals.[21]

This 'naturalness', however, is just one of several issues raised in the account. Wierzbicka begins, for example, by presenting arguments against what she calls the 'Ekmanian paradigm', briefly discussed in the previous section. She claims that the time is ripe for a new direction in the study of human facial expression: 'A fresh breeze is blowing in the field...and there is a widespread

[21] Though this is not to suggest they might not somehow *become* coded.

sense that the time has come for a change of paradigm' (2000, p. 147). She lists ten basic assumptions fundamental to this new direction.[22] These assumptions lay the foundations for what is essentially her main claim, that the coded element of human facial expressions can be analysed in terms of the same 'Natural Semantic Metalanguage' we saw in the previous chapter; below is another example:

[22] I have chosen not to list all ten assumptions here, intending as I do to focus on Wierzbicka's Natural Semantic Metalanguage (NSM). However, I would briefly like to comment on two: (a) 'a semantic analysis of the human face...requires the identification of minimal meaningful units of facial behaviour'; (b) 'we need to distinguish the "semantics of human faces" from the "psychology of human faces"' (2000, p. 150). Regarding (a), rather than analysing 'eyebrow flashes', 'smiles' or 'frowns' Wierzbicka's proposal is that we should be analysing *moving one's eyebrows upwards, doing something with one's mouth in such a way as the corners of one's mouth move upwards, moving one's eyebrows so that they will be (relatively) close together...* It is claimed that these units can shed more light on the meaning of facial expression than, for example, Ekman's own sophisticated 'Facial Action Coding System' (FACS) (Ekman and Rosenberg 1997): 'if we are interested in *meaning* we must adopt the perspective of "the ordinary people" who want to communicate with one another, and not that of a physicist working in a laboratory' (Wierzbicka 2000, p. 159). I think Hauser's point about two different levels of analysis (see pp. 192-193 above) is relevant here. FACS is a highly sophisticated, rigorous system for measuring facial movements in great detail: it is neither concerned with the recognition of facial expressions, nor directly with their communicative content. That being said, one alternative to Wierzbicka's proposal that the FACS should be replaced with her 'minimally meaningful units' (and one which I think I favour) would be to try and *integrate* Ekman's exhaustive work on facial expression into an account of their 'meaning'. It may, after all, prove more fruitful in the long term to build bridges between the two disciplines. This point carries over to assumption (b). Surely a plausible theoretical account of the semantics of facial expressions will *mesh* with theories of the psychology of facial expressions, not replace them.

(7) courageous

X is courageous. =

X can do very good things when other people can't

because when other people think something like this:

I don't want bad things to happen to me

X thinks something like this:

it is good if I do this

it is bad if I don't do it

I want to do it because of this

this is good

Having extended the approach to the meaning encoded by interjections, Wierzbicka now applies it to facial expressions as in (8) below:

(8) *raising of the eyebrows*

I know something now

I want to know more (about this)

I'm thinking now

In the previous chapter I outlined what I regard as major problems with the conceptualist approach to interjections; the same arguments carry over to the proposed analysis of facial expressions. However, of the numerous issues that Wierzbicka's account raises, I would like to focus primarily on three problems that I see with the approach: the first is a more general one, concerning the model of communication it presupposes; the second is a (related) more specific problem

concerning the tools of the analysis itself; the third concerns a specific area of confusion that seems to be present in the account.

The first, general problem echoes a point I have made at various places in this thesis against the semiotic program generally. The conceptualist approach to human communication is highly reminiscent of just the kind of coding-decoding model which I am rejecting. Wierzbicka (1996, p. 8) talks in terms of 'the meanings encoded by natural language', and largely ignores the kind of insights that have been discussed concerning the link between intention and meaning.

Consider the following on facial expressions: 'Human faces send messages, and these messages *must be decodable*' (2000, p. 178—my emphasis, TW). Whilst I agree that there may be a coded element to some facial expressions, I do not agree that for a facial expression to communicate something it necessarily has to encode *anything at all*. If I catch your eye during a boring presentation and look ostensively toward the door, I might communicate to you that I want to leave, but it is not obvious that communication is achieved by my encoding anything. If I deliberately and openly let you see my spontaneous shiver, I am not encoding the conceptual structure 'I feel cold', any more than if I point to a cloud I am encoding the conceptual structure 'it's going to rain'. Information does not have to be *encoded* to be successfully *communicated*, and this is overlooked by NSM-based accounts. What is required are clear criteria by which we can decide whether a behaviour is coded or not (we have seen that evolutionary function is the key motivation, as in Hauser's distinction between signs and signals), and if there *is* a coded element to a particular facial expression, then this needs to be set within an inferential framework.

The second, more specific, disagreement with Wierzbicka's account relates to the conceptual structures themselves, and how adequately they capture natural codes. Facial expressions are marvellously versatile, and Wierzbicka is careful to include among her fundamental premises the assumption that 'semantic analyses (whether of verbal utterances or facial expressions) must distinguish between the context-independent invariant and its contextual interpretations' (2000, p. 151). There is a semantics/pragmatics distinction in conceptualist analyses, and she is clear to point out that what is being characterised by these conceptual structures is *what is encoded*, rather than *what is communicated.*

Consider the eyebrow flash (or *raising the eyebrows*). This is one candidate for a universal facial expression[23] (though see Ekman 1999) that has been much discussed. In a seminal paper, Eibl-Eibesfeldt (1972, p. 300) describes some of its various functions as follows:

> 'We mentioned several situations in which eyebrow flashes of approximately the same stereotyped form occur: greeting, flirting, approving (yes), seeking (asking) confirmation, thanking and emphasising a statement (calling for attention)... Finally, we raise the eyebrows during disapproval, indignation, and when we look at a person in an admonishing way'.

Eibl-Eibesfeldt concludes that if there is a common denominator in all these functions it is that the eyebrow flash is a kind of 'attention' signal: 'the basic common denominator is a "yes" to social contact, and it is used either for requesting such a contact or for approving a request for such contact' (1972, p. 300). Wierzbicka comments that her structure is not inconsistent with

[23] And to my mind a prime candidate for a natural code.

interpretations relying on the technical expression 'attentional activity' (2000, p. 168).[24]

But it is hard to see how. The gap between the conceptual structure in (7) and the uses described above is so vast, that it seems implausible to suggest that they are all pragmatically derived variants of this context-independent structure. Furthermore, these rigid conceptual structures do not begin to capture the 'natural', i.e. 'iconic or indexical basis' Wierzbicka (I think quite rightly) observes.[25] Conceptualist structures are entirely digital constructs and as we have seen, in the case of human natural codes, what is needed is some way of accounting for their *analogicity*.

Interestingly, although Wierzbicka remarks on the natural—or analogue/indexical—side of facial expressions, she chooses *not* to contrast it with the digital nature of language. On the contrary, one of her fundamental assumptions is that we should stress the similarities between the two: 'facial expressions can convey meaning comparable to verbal utterances' (2000, p. 151). This brings me to the third problem. In a sense, of course, it is true that the meanings conveyed by facial expressions are comparable to those conveyed by verbal utterances. You give me a gift, and I take it from you, smiling broadly. The meaning I convey to you might be paraphrased as 'Tim is delighted with the gift'. If I choose to convey this information to you by saying 'I am delighted with

[24] Though she does not endorse the use of the expression. Wierzbicka is actually referring to a paragraph from Smith and Scott (1997, p. 239), and takes issue with their terminology: 'The problem with this approach is that it is not quite clear what precisely is meant by "attentional activity", and since this expression does not belong to ordinary language we can't use our ordinary linguistic intuitions to interpret what exactly the writers really have in mind' (2000, p. 164).

[25] Allowing for my previous reservations over iconic representations being coded.

the gift', the meaning conveyed is certainly comparable. Notice, however, that it is *only* comparable if we take 'the meaning conveyed' by verbal utterances and facial expressions to be what they *communicate*, rather than what they *encode* (unless smiles and language work in exactly the same way—and if they do, then we may as well say that when I point at a cloud, I *am* encoding 'it's going to rain').

This element of confusion is unavoidable within the conceptualist framework. The root of the problem lies in the code model foundations upon which it is built. Inference is relegated to a minor role, an 'add-on' to a human communicative process that is fundamentally characterised in terms of coding and decoding. This leads to the assumption that pretty much the only way to communicate a concept is to encode it, which in turn leads to a position on the relationship between language and thought that I also believe to be problematic.

One problem that Wierzbicka raises with work within the 'Ekmanian paradigm' is that Ekman's use of English words to label the emotions signalled by universal facial expressions results in an 'ethnocentric' view of their interpretation: since the English word 'anger' differs in meaning to the Italian word 'rabbia', the former cannot be said to be a 'universal category of human experience' (1994, p. 439). Actually, it is clear from Ekman's work that his claim that universal facial expressions reflect the existence of universal emotions does not in any way presuppose a one-to-one mapping between these emotions and the words speakers of different languages might use to label them: in fact, he is quite explicit on this point: 'we never claimed that facial expressions evolved to

represent specific verbal labels. Nor did we say that the meaning of an expression is limited to or best captured by a single, specific word' (1994, p. 270).[26]

However, what these criticisms also reveal is a Whorfian, relativistic view of the relationship between language and thought: 'speakers of other languages...think about human experience in terms of other non-matching conceptual categories...they do not read human faces as "angry"...but rather interpret them in terms of their own language-specific categories' (2000, p. 149). Thought takes place in words, and is shaped by language. Under such a view, it follows that most concepts are lexicalised. If they were not, the efficiency of thought processes, as well as communication, would be seriously compromised. Unlexicalised concepts would not be entertainable, let alone communicable.

A consequence of the relevance-theoretic inferential view of communication is that it is not the case that most concepts are lexicalised. Individuals are capable of acquiring an enormous amount of information each day, and it seems implausible to suggest that items of information can only be stored in the mind if there are public words which encode this information. Sperber and Wilson argue that, in fact, 'there are many times more concepts in our minds than words in our language' (1998, p. 196), and that in everyday thought we regularly entertain unlexicalised concepts. We communicate them regularly too, for as we have seen a concept that is lexicalised may be inferentially enriched by a hearer to derive

[26] Wierzbicka's response to this quote from Ekman is to quote him again, this time from his rebuttal of Russell's critique (1994, p. 276): 'Russell complained that we and others had pre-selected our expressions [i.e. emotion labels, A.W.]'. However, the word 'expressions' in this quote does *not* refer to 'emotion labels' as Wierzbicka indicates, but instead to photographs of the facial expressions themselves. (This is not to deny that Russell does in fact take issue with the 'forced-choice format' of Ekman's experiments—see Ekman (1994, pp. 273-275) for his response.)

the (slightly or substantially different) sense that a speaker intended to convey. This represents another argument against the position taken in NSM analyses, for it suggests that not only do humans engage in non-verbal, non-coded communication, but also that words routinely communicate *ad hoc* concepts which differ from the precise conceptual content they encode (see discussion of narrowing and loosening in Chapter Three).

As it stands, I think these three arguments suggest that conceptualist accounts of facial expressions are at least implausible: as it is in semiotic approaches generally, communication is seen largely as a coding-decoding affair; the conceptual structures they propose do not reflect the *analogicity* of human natural codes; the relationship adopted between language and thought leads to inevitable confusion concerning what it actually is that the conceptual structures are designed to characterise. As yet, however, other than arguing that natural codes should be integrated within an inferential model, I have offered no alternative to the conceptually encoded structures that Wierzbicka proposes. Does such an alternative exist? We saw in the last chapter that it does: the information encoded in natural codes is not *conceptual* at all.

2.2 Non-translational encoding II

According to the account of interjections offered in the previous chapter, the information encoded by an interjection merely *encourages* the hearer to embed the proposition expressed under speech-act or propositional-attitude descriptions by constructing higher-level explicatures. In some cases, as we saw, there is no propositional embedding and use of an interjection is directed at a suddenly

perceived object or event. The information encoded by the interjection is *non-translational*—in relevance-theoretic terms, *procedural.*

Whether this non-translational account of interjections really does amount to a procedural one remains to be seen. Firstly, the very nature of the kind of information encoded by non-translational encoding is somewhat different to the kind of information portrayed as procedural in earlier accounts (Blakemore 1987). Indeed, although in Chapter Three I presented my translational/non-translational distinction as a way of *introducing* the conceptual-procedural distinction, it was originally conceived of as a way of *broadening* the notion of procedural meaning (see Wharton 2001).

Recall from Chapter Three that in much work on the distinction, procedural information is characterised as *instructions* to the hearer. The notion of procedural information as (potentially) vague indicator adopted here contrasts with this view. It is my belief, however, that the broader view is worth exploring; moreover, a broader notion of procedural meaning is required anyway in order to accommodate the full spectrum of linguistic devices currently seen as encoding procedural information (see discussion in Chapter Three, pp. 130-131).

Secondly, procedural analyses have only so far been proposed for properly linguistic items and, as I remarked earlier, the linguistic status of interjections is marginal at best. If interjections are not part of language—and I believe that in the previous chapter I have suggested good reasons why they are not—then this would suggest a degree of dissociation between procedural and linguistic encoding. Notice, however, that much depends on whether we take the notions of procedural and non-translational encoding to be co-extensive, or whether we regard procedural encoding as a sub-type of non-translational encoding.

Notwithstanding what are essentially theory-internal considerations, we shall see below that *non-translational* encoding is not linked exclusively to linguistic expressions: interjections, then, can be seen as encoding non-translational information *whether or not* they are part of language.

The implications for an analysis of natural codes are clear. True, these signals are not part of the *linguistic* code, but they are coded signals, and they too are best analysed along non-translational lines. Consider examples (9ab):

(9a) Mary: (in a regretful tone of voice) I don't feel well.

(9b) Jack: (smiling happily, in a pleased tone of voice) John has arrived.

Hearers of these two utterances would be led to form the conceptual representations—in relevance-theoretic terms, higher-level explicatures—in (10ab) below:

(10a) MARY REGRETS [THAT SHE DOESN'T FEEL WELL]

(10b) JACK IS HAPPY [THAT JOHN HAS ARRIVED]

Notice that the higher-level conceptual representation embedding the basic propositional content is activated non-translationally, and in both examples the route to the conceptual representation the hearer entertains patterns with the earlier analysis of interjections ((10b), for example, is almost entirely parallel to example (24b) from the previous chapter). The proposal, then, is that the coded element in all manner of natural behaviours—Mary's tone of voice in (9a),

Jack's smile in (9b), vocal and facial gestures generally—are best analysed as encoding non-translational information.

There are a variety of arguments to support this claim. Like interjections (and *linguistic* expressions that are analysed as encoding procedural information), facial expressions do not contribute to the truth conditions of utterances. If Jack says to Lily 'I am happy', she might reply 'That's not true, you aren't happy'. If he simply looks at her and smiles, she would be unlikely to make the same accusation. The same arguments can be put forward for crying and shivering (or any other non-verbal behaviour for that matter).[27]

Nor do facial expressions compose to form larger 'phrases'. Of course, it is true that smiles, eyebrow flashes, frowns and gestures are 'discrete' signals insofar as they are 'distinct'. However, they are not digital in the sense that they combine to form different meanings. This point is easily demonstrated if you compare the putative discreteness of the natural manual gestures that accompany speech, for example, with the genuinely compositional component 'gestures' of sign-language proper, which are true digital systems. This observation is found in the work of Adam Kendon (1988), who sees gesture generally as existing along a continuum from *gesticulation*—the spontaneous movements that accompany speech, through *pantomimes* and *emblems*—culturally-regulated gestures, to signing proper. It is also reflected in Ekman's (1999) own distinctions among non-verbal signals.

[27] Nods and headshakes may be an exception. If I ask you if you *really* don't want another glass of wine, and in response you nod, I think I may well be justified in saying 'That's not true; you *do* want one'. On a different note, see Darwin (1872) for a proposal that behaviours such as head shakes have their roots in entirely natural behaviours, perhaps along the lines of ethological intention movements (Darwin's proposals are also discussed in Eibl-Eibesfeldt (1972, p. 305)).

The communicative content of facial expressions is vague and context-dependent too. As with interjections, this is captured by proposing that what is activated (non-translationally) by the coded element of the facial expression is a cognitive state that might include the (non-translational) triggering of a variety of emotion or attitudinal concepts, which then further constrain inferential processes.

And finally, the non-translational quality of the information encoded in natural codes provides the key to capturing their analogicity. Natural codes do not encode digital conceptual structures, but rather point the audience in the appropriate direction. The mental state that is activated is activated to a degree that is consonant with the breadth of the smile, or gravity of the frown. Of course, unlike the engineer assessing the pressure in the system, we do not read the quivering needle consciously. As discussed earlier, it is an automatic decoding process, rather like the honeybee calculating the distance between the hive and the nectar according to the intensity of the communicating honeybee's dance.

So *what* do natural behaviours indicate? An ostensive stimulus is typically a composite of natural (and semi-natural) as well as linguistic signals. When accompanying linguistic acts, these behaviours indicate information about the speaker's intended meaning, and contribute to the construction of higher-level representations—in relevance theoretic terms higher-level explicatures. Typically, this information is attitudinal or emotional.

This question is harder to answer in cases where a natural behaviour is used *alone* as an ostensive stimulus. In the last chapter I suggested that in such cases interjections—as is the case with non-verbal communication generally—

communicate implicitly. I also mentioned that I thought there was something inherently problematic in this, and am unhappy to carry this over to my analysis of natural codes. Not only does it seem wrong to call naturally coded information *implied* information, but within relevance theory, implicit content is, by definition, wholly derived via inference.

One solution would be to see natural codes, even when used alone, as communicating explicitly. Another proposal would be to introduce an explicit/implicit distinction for natural behaviours; thus, we might speak of the explicit natural content, and the implicit natural content of a communicative act. However, this proposal also requires further thought and development, not least because in the current relevance-theoretic picture, ostensively-used paralinguistic behaviours encourage the formation of higher-level explicatures *whether or not* they are coded signals. One suggestion would be that all natural behaviours, when used alongside linguistic utterances, encourage the formation of higher-level explicatures; the natural 'explicit-implicit' distinction only comes into play when such behaviours are used alone.

With regard to how natural behaviours are interpreted, those natural behaviours that have the adaptive function of indicating are best analysed as natural codes. On one level, they work in a similar way to animal communication systems. However, among humans they have acquired an intentional communicative function. This is not surprising, given that humans are aware of their involuntary responses in a way that animals are not (see Allen and Bekoff 1997 for discussion). As we saw in Chapter One, natural behaviours are routinely deliberately shown to provide the audience with clues as to the communicator's informative intention. The fact that some natural behaviours are coded signals

would predict that they are interpreted by specialised, perhaps dedicated, neural machinery. This prediction appears to be borne out. Both non-human primates and humans have neural mechanisms dedicated to both recognising faces and processing facial expressions (Gazzaniga and Smiley 1991).

The term 'natural' covers a wide range of behaviours, about which the distinctions presented here allow us to be more precise. An utterance may sometimes, for example, be accompanied by ostensively used *non*-coded behaviours—a spontaneous shiver, for example. This may well be picked out by the relevance-based comprehension procedure. It may also be accompanied by other (non-ostensive) natural behaviours. These other natural behaviours are inherently *non*-communicative (signs, as opposed to signals). Assuming they are noticed, these signs will still be processed by the audience, but, crucially, not in order to form a hypothesis about the communicator's *intended* meaning. For if a natural behaviour is not used ostensively, or recognised as ostensive, it will not be picked up by the relevance based comprehension procedure.[28]

The information encoded in these natural codes is non-translational in the sense described above. It is information that activates particular internal states which are *analog*ous to the intensity of the signal. The decoding process is automatic, and leads to the non-translational activation of a range of attitudinal and emotion concepts, which help constrain the inferential search for relevance. For the fact that a communicator has made it mutually manifest that there is an informative intention behind the deliberately shown behaviour means that the automatic decoding processes will be supplemented by other processes—perhaps equally specialised—that govern the search for relevance.

[28] Assuming it can make such subtle distinctions. There might, for example, be a situation where the procedure is triggered 'accidentally', as it were.

There is a whole range of behaviours that encode this kind of information, ranging from linguistic devices such as discourse connectives, pronouns and mood indicators on the one side, to interjections and coded facial expressions on the other; the approach might also be carried over to some of the prosodic aspects of speech.

Human speech, of course, exhibits a wide range of such features; speakers use stress, rhythm and pitch change in varying degrees to help convey their intended meaning. In many respects these too seem to exist at various points along the showing-meaning$_{NN}$ continuum. In languages such as Thai and Burmese, for example, pitch change is phonologically contrastive, and can be used to convey linguistically-encoded lexical meaning. Even in languages with no lexical tone, intonation can indicate interrogative mood. Far at the other extreme of the continuum—the showing end—we find affective tone of voice; so Jack might infer that Lily is happy or sad, relaxed or tense, engaged or bored solely on the basis of the voice quality. We might think of other suprasegmental features of English, such as sentence stress and intonation as falling at various points between the two extremes.

Regarding intonation in particular, Dwight Bolinger (1983) is of the view that we would be better to focus more on its natural side. Indeed his analysis might be seen as an attempt to characterise it as one kind of natural code, focussing as he does on the interaction between intonation, and other 'natural' components of the complex communicative stimulus:

> 'If intonation is part of a gestural complex whose primitive and still surviving function is—however elaborated and refined—the signalling of emotions and their degrees of intensity, then there

should be many obvious ways in which visible and audible gesture are coupled to produce similar and reinforcing effects. This kind of working parallel is easiest to demonstrate with exclamations. An *ah!* of surprise, with a high fall in pitch, is paralleled by a high fall on the part of the eyebrows ... A similar coupling of pitch and head movement can be seen in the normal production of a conciliatory and acquiescent utterance such as "I will" with the accent at the lowest pitch—we call this a *bow* when it involves the head, but the intonation bows at the same time.'

(1983, p. 98)

He is, however, keen to stress that behaviours may be *more or less* natural (or non-natural) and that even though we may feel some aspects of intonation to be grammatical, there still a sense in which they have their roots in natural behaviours:

'Intonation ... assists grammar—in some instances may be indispensable to it—but it is not ultimately grammatical ... If here and there it has entered the realm of the arbitrary, it has taken the precaution of blazing a trail back to where it came from.'

(1983, pp. 106-108).

If, as Bolinger appears to suggest, there *is* a diachronic dimension to the continuum between display and language, then it may turn out to be a useful tool with which to follow the trail back from arbitrary expressions to their natural origins.

In a recent paper, Carlos Gussenhoven (2002) presents evidence which suggests that in the case of intonation, Bolinger may be right. Some aspects of intonation, he argues, are arbitrary and properly linguistic. Other, universal,

aspects of intonation, however, are the result of the existence of 'biological codes' which correlate high and low pitch, the presence or absence of articulatory effort and other aspects of the production process with what he describes as 'universal paralinguistic meanings' (2002, p. 47). Gussenhoven also notes that the linguistic aspects of intonation are digital, while the 'paralinguistic' aspects are analogue. This represents exciting evidence concerning the existence of natural prosodic codes.[29]

[29] For reasons discussed in the introduction, I prefer 'natural code' to 'biological code' (language is a biological code too).

Chapter Six

1. Mindreaders

'The words "know" and "feel" were like "it" and "of" and "by"—you couldn't see them or touch them, so the meaning wasn't significant. People cannot show you a "know" and you cannot see what a "feel" looks like.'
(Donna Williams—*Somebody Somewhere*)

According to the account of communication adopted throughout this thesis—from Grice's earliest work through to relevance theory—linguistic communication is an intelligent, intentional, inferential activity. Utterances do not encode the messages speakers want to convey; rather, they are used to provide evidence of intentions, which hearers must infer. Although the extent of the role played by inference in communication is still the subject of much debate (as indeed is the precise nature of what constitutes 'inference' itself)[1] most pragmatists now agree that verbal communication is more than a simple coding-decoding process.

It's worth remembering, however, that it is not just when involved in acts of communication that humans are inclined to attribute mental states to others. The human disposition to attribute mental states is so much a part of our individual (or collective, species-specific) psychological make-up, that—as I know to my cost—it is not something we can choose to do or not to do: it's something *we just can't help*, any more than we can help pulling our hand back from a source of extreme heat.

Plainly, other people's intentions and mental states generally are not objects to be perceived in the world in the same way as are their faces or bodies; they are

[1] See Recanati 2002.

'out there', but they are invisible. It is hard, however, to even imagine what it would be like not to be able to sense the mental states of others in some way. The world would be such a different, potentially terrifying place. The human thumb accounts for over 50% of the function of the human hand; we can touch it, we can see it and we can feel it. Yet it is still very difficult to imagine how we might cope without one. In the case of our thumb, of course, we are given a salutary reminder each time we injure it—just try tying your shoelace, or riding a bike with a sprained thumb. In the case of what it would like to be unable to attribute mental states we are left with thought experiments of the kind suggested by Baron-Cohen (1992, pp. 1-5) and the few first-hand accounts of what it is really like. Indeed, the central role the recognition of intentions plays in human interaction generally is no more clearly illustrated than by the accounts of those individuals for whom the mental states of others, rather than being merely out of sight, are locked away—permanently out of reach (Williams 1992, 1994, 1999; Holliday Willey 1999; see also Happé 1992 on the autobiographical writings of three Asperger syndrome adults, and Sacks 1994).

The author of the above epigraph—Donna Williams—is autistic. Her autobiographical works 'Nobody Nowhere', 'Somebody Somewhere' and 'Like Colour to the Blind' are vivid accounts of what it is like to be *mindblind*—to use the term adopted by Simon Baron-Cohen (1995). Donna's world is a strange, unfamiliar, frightening one: a world of 'inner isolation', of 'persistent aloneness' (1994, p. 95); a world it took enormous strength and courage to escape.

Our understanding of autism is still limited; while it is not in its infancy, research is certainly at an early stage of development. But great strides have been made, and there is a growing literature on both the precise nature of the deficits

and impairments that give rise to the condition, and the effects autism has on the capacity of autistic people to communicate and interact with other people (Leslie 1987, Happé 1994, Scholl and Leslie 1999).

Baron-Cohen (1995) makes some concrete proposals about the specific deficits that might lie behind the condition, and suggests that it is characterised by a (partial) breakdown in the mechanisms underlying the human mind-reading ability. The system breaks down into four sub-components—an *Intentionality Detector*, an *Eye-Direction Detector*, a *Shared-Attention Mechanism* and a *Theory of Mind Mechanism*[2]—the existence of which is reflected in various stages of the psychological development of normal children, and illustrated by instances of pathological breakdown (as, for example, in the case of autism).

Baron-Cohen's hypothesis is that autistic subjects exhibit a deficit in their Shared Attention Mechanism.[3] This has two knock-on effects: firstly, it follows that they cannot construct complex three-place relations, such as '*He sees (that)*

---

[2] The *Intentionality Detector* is responsible for the interpretation of the movement of 'agent-like' objects in terms of basic volitional concepts such as *goal* and *desire*; the *Eye-Direction Detector* (fairly transparently) detects 'eyes' and, more importantly, monitors them in order to decide whether they are directed toward *it*, or is directed toward some other object or organism; given input from the Intentionality Detector and the Eye-Direction Detector, the *Shared-Attention Mechanism* allows individuals to monitor which objects, events and states they and another individual are jointly attending to. (In relevance-theoretic terms, it keeps track of what is mutually manifest.); the *Theory of Mind Mechanism* allows the individual to infer not only primitive volitional states (such as 'goal' and 'desire'), and perceptual mental states (such as 'see') but also what Baron-Cohen calls 'epistemic mental states' (1992, p. 51), such as 'believing' and 'thinking'. Baron-Cohen's Theory of Mind Mechanism is based on the one presented in Leslie (1994).

[3] There is at least some evidence from Donna Williams' accounts that she may have some deficit with her *Intentionality Detector* too. Consider the following: 'Everything had its own, if limited, volition. Whether a thing was stationary or movable depended more on the thing's readiness to move than on the person's decision to move it. Statements like "It won't budge" only confirmed this assumed reality' (1994, p. 65).

[I see her]', and as a result cannot grasp that they and another person are attending to the same object; secondly, and more crucially, there is no output from the Shared Attention Mechanism to trigger the development of Theory of Mind Mechanism, which, it is claimed, is the mechanism underlying the human ability to attribute complex epistemic mental states (or propositional attitudes) such as 'believe' and 'think'.

Imagine two people (A and B) running past you quickly down a road, one behind the other. Rather than viewing their motion as you might view the random motion of two billiard balls—one apparently 'following' the other around a billiard table—it is an automatic response to interpret A and B's behaviour in terms of the mental states that underlie it. At a basic, volitional level, you would perceive their actions in terms of their *goals* and *desires*; in terms of the person behind—A—wanting to catch the second—B, or B's desire to escape.[4] Such notions, according to Baron-Cohen (1995), are 'primitive mental states in that

[4] You would almost certainly interpret a dog chasing a cat in terms of their respective goals and desires too, but like me you may find yourself less willing to attribute to either the cat or the dog more complex, epistemic mental states of the kind discussed below. (For discussion see Allen and Bekoff 1994, Chapters 2 and 6). Among undergraduates—and some ex-Heads of Department—the behaviour of household pets is a commonly cited source of evidence used as a counter-argument to claims concerning the uniqueness of the human mind-reading ability—'But my dog knows *precisely* what I'm thinking', the argument goes.)

Baron-Cohen stresses that the mechanism responsible for detecting 'agency' in general would also be triggered by agent-*like* objects, and thus prone to false positives (he argues this would be an evolutionary advantage). So, strangely enough, the tendency may well be to interpret two billiard balls apparently 'following' one another around a billiard table in a manner similar to the way in which you would interpret the two people running in my example, rather than the other way round. On realising that the motion of the billiard balls was not self-caused, the thinking goes, your original reading of the situation would be revised.

they are the basic ones that are needed to be able to make sense of the universal movements of all animals: approach and avoidance' (pp. 32-33).[5]

However, as well as these basic concepts you would also attribute to A and B more complex, epistemic states. So, for example, it might be clear to you that A *believes* that B has done something wrong (the complex mental state underlying A's goal), and that B believes that A intends to catch him (the complex mental state underlying B's desire to get away). Notice that the kind of beliefs you attribute may well be incompatible with the beliefs you yourself hold: you have no problem attributing to A the belief that B has stolen his wallet, even though you know that B has not (because, for example, you have A's wallet in your own pocket). Autistic people typically fail tests based on this kind of 'first order' theory of mind ability, and this is generally taken to be evidence that they are unable to attribute more complex, epistemic mental states (Dennett 1978).

Since the mind-reading ability is recursive, our own and others' mental states can be *meta*represented in a multi-layered way, to a number of different levels. Watching the chase, you have no problem entertaining the thought that B believes that A knows that he has done something wrong, or even that A believes that B believes that A knows that he has done something wrong (or even, for that matter, that B knows that A believes that B believes that A knows that he has done something wrong).[6]

---

[5] Given his hypothesis, many autistic subjects would indeed be able to interpret the situation described above in terms of the goals and the desires of the individuals concerned; Baron-Cohen notes, however, that 'this does not mean that they are able to understand all aspects of desire, or the more complex mental state of intention' (1995, p. 63).

[6] Dennett (1988, pp.185-186) wonders about the upper limits of the human recursive, meta-psychological ability: 'How high can human beings go? "In principle", forever, no doubt, but in fact I suspect that you wonder whether I realise how hard it is for you to be sure that you

It is a recursive, metapsychological ability such as this that is exploited in cases of intentional communication. At this point, therefore, the border between Grice's philosophy, modern cognitive science, psychology and (even) cognitive ethology becomes so blurred as to disappear. There is thus a point of contact (though, of course, agendas may vary) between the literature discussed in Chapter One, intention-based pragmatic models that Grice's work has inspired and recent psychological research on the capacity among humans and non-human animals to attribute mental states to one another. It is this point of contact between disciplines that this work (and, I hope, future work) explores.

It seems clear that the kind of meta-communicative abilities necessary for intentional communication in Grice's sense and ostensive-inferential communication as defined in relevance theory are related to the wider meta-psychological mind-reading ability discussed above. On the strength of *A*'s utterance of 'Stop! Thief!' and the accompanying glance at you, you might attribute to him the intention to get you to believe that *B* is a thief too (or at least the intention to get you to think that *he* believes *B* is a thief); furthermore, that he wants you to help. By your hurried glance back, you might communicate to him that you do indeed intend to help.

There are good reasons to suggest, however, that there is more to the interpretive processes that underlie *verbal* comprehension (or, more generally, ostensive-inferential communication) than the kind of general mind-reading abilities responsible for our interpretation of A's and B's actions in my example above. Firstly, the content of a speaker's intention—the 'meaning' she intends to convey—will invariably be more complex than the kind of intention you might

---

understand whether I mean to be saying that you can recognise that I can believe you to want me to explain that most of us can keep track of only about five or six orders…'.

attribute to someone on the basis of observable evidence such as the chase between A and B in the example sketched above. Secondly, it should not be overlooked that one of the reasons that you are able to attribute to A the intention to catch B is that you can *observe* his efforts: by contrast, in the case of communicators' intentions the only real clue an audience has as to the content of the complex intention which constitutes the communicator's meaning, is the fact that a communicator has provided evidence that she has one: there is, on the face of it, nothing more to go on.[7] Thirdly, and perhaps most crucially, in cases of intentional communication, it is always the case that *several* layers of metarepresentations are necessary; yet young children below the age of 4—the same children who (as do autistic subjects) regularly fail basic 'first order' theory of mind tests—master verbal communication quickly and effortlessly well before this age.

These and other arguments (see Sperber and Wilson 2002, Wilson and Sperber 2002) have led to the proposal that the processes that underlie verbal comprehension might be performed by a specialised, domain-specific 'comprehension' mechanism or *module*[8] (Sperber 1996, 2000). The task of such a mechanism would be to interpret ostensive stimuli according to the relevance-based comprehension procedure outlined in Chapter Three. Such a sub-module may form part of the wider mind-reading ability (in the same way as Baron-Cohen's 'Shared Attention Mechanism').

---

[7] The solution, of course, is the Communicative Principle of Relevance (and the relevance-theoretic comprehension procedure it motivates) .

[8] I use the term 'module' in the sense of Sperber 1996 and, indeed, much of the literature on evolutionary psychology: that is, in a somewhat 'looser' sense than the one originally proposed by Fodor 1983.

Indeed, despite the fact that young children below the age of 4 regularly fail basic 'first order' theory of mind tests, there is some experimental evidence which suggests that they are able to track false beliefs in word-learning tasks *before* they can pass false belief tests (Happé and Loth 2002). This further supports the proposal that the mind-reading abilities which are a prerequisite for verbal communication dissociate to some degree from the wider human mind-reading ability; it also supports the hypothesis that there is a separate, comprehension module.

There is a growing literature on the question of how it is that children can be adept interpreters of utterances before they can read minds (in the sense of pass regular false-belief tasks). Broadly speaking, two camps are emerging: on the one side, there are those (like Happé and Loth—see also Tomasello and Barton 1994, Akhtar, Carpenter and Tomasello 1996, Bloom and German 2000) who warn against underestimating the cognitive prowess of young children; on the other, there are those who claim that we *over*estimate the degree to which the inferential attribution of intentions is a prerequisite to verbal communication. Two recent papers from protagonists of the latter camp are Breheny (forthcoming) and Recanati (2002).

Breheny proposes an account of 'basic' communication,[9] the central claim being that it does *not* require the attribution of propositional-attitude mental states (or presuppose theory of mind). The account exploits aspects of relevance theory, in particular the relevance-theoretic notion of mutual manifestness, together with the notion of *joint attention to a shared situation*, as originally proposed in Barwise 1989. Breheny agrees with Sperber and Wilson that act of

[9] A communicative act typified by the assertive utterance of a declarative sentence.

basic communication 'involves one agent drawing another agent's attention to something' (forthcoming, p. 46), but, according to him, the objects of such acts are not mental states but situations (in a shared situation).[10]

There are, however, a number of issues which suggest the account may be overly minimalistic. In the first place, children are clearly attributing mental states to others well before they pass the 'false-belief' task, perhaps as early as one year old (see Bretherton 1991). Leslie (1987) claims that children come to understand (what he refers to as) the *epistemic* mental state 'pretend' between the ages of one and two years. More generally, very young children respond effortlessly and appropriately to *non*-verbal cues that indicate the mental state of a parent, carer or friend. If we put this together with the fact that children develop the shared attention ability between nine and eighteen months of age, it seems an overly minimal proposal to suggest that the objects of acts of basic communication cannot be mental states of some kind, or involve mental states in some way.

Secondly, as we saw in the discussion in Chapters One and Three, even in the most basic acts of ostensive inferential communication,[11] cases in which a

[10] No relevance-theorist would argue against Breheny's claim that communicators point to objects, situations etc., and in that sense his view is perfectly compatible with the relevance-theoretic viewpoint. Where the two approaches diverge is that Breheny goes on from this observation to deny that the attribution of mental states has a role to play in basic communication, whereas—as I discussed in Chapter Three, and reiterate below—on the relevance-theoretic view, the mental states of the shower or pointer invariably play a crucial role in determining: (a) *what* it is that has been shown or pointed out; (b) *why* it has been shown or pointed out in the first place.

[11] Breheny does not provide a definition of 'ostensive-inferential communication' since, as an everyday concept, 'communication' cannot be defined (in this he is following the methodological strategies of Fodor 1998 and Chomsky 2001): 'a communicative situation is that kind of communication which our concept of communication locks onto as a result of exposure to

communicator provides *highly* direct evidence of an intention to inform (e.g. my mad photograph-shower example) the mental states of the shower still play a role in correctly determining what it is that has been shown, and why. Evidence suggests that children begin understanding pointing gestures at around fourteen months (Blake, McConnell, Horton and Benson 1992); according to Leslie (1994): 'informative *showing* typically makes its appearance early in the second year along with verbal communication'. It is not clear how this can be done solely on the basis of the mutual manifestness of a situation (in a shared situation), and *without* reference—to at least some degree—to the mental states of the shower.

Finally (albeit anecodotally) as the father of two young children myself, I find myself constantly amazed by their communicative (and cognitive) precocity, rather than frustrated with their inability.[12] It just seems unwise to underestimate them; the jury, however, is still out.

While Recanati (2002) does not specifically discuss theory of mind, his assertion that relevance theorists overestimate the role played by inference in verbal communication might also be said to put him in the second of the two camps described above. According to Recanati, communication is not constitutively inferential; instead, it is—at the explicit level at least[13]—'as direct as perception' (2002, p. 105). Mature communicators may indeed use indirect

---

stereotypical instances of communication' (ibid., p. 51). Surely, though, as theorists, we should specify what domain our intended generalizations are intended to apply to

[12] Of course, it may be that I am simply guilty of projecting onto my children abilities they do not have in a manner analogous to the individual in *fn.* 3 above, who projects abilities onto his/her household pet.

[13] Recanati does not use the notion of explicit content, preferring his own pragmatically enriched notion of *what is said.*

means in order to convey their meanings—implicate them, for example—but the ability to infer speaker meaning is not a necessary condition of being able to communicate verbally. While relevance theorists stress the role of pragmatic inference in the derivation of the explicit content of an utterance, Recanati proposes that while these processes are indeed pragmatic, they are not inferential. As he puts it:

> '[T]hose pragmatic processes that are involved in the determination of truth-conditional content—*primary* pragmatic processes, in my terminology—need not involve an inference from premises concerning what the speaker can possibly intend by his utterance. *Indeed, they need not involve any inference at all.*'
>
> (2002, pp. 113-114)

Ultimately, a great deal will depend on how we define inference. Indeed, Recanati's arguments rest on his utilisation of a notion of inference that is distinct from the one adopted in relevance theory, a full discussion of which would take us too far afield here.[14]

Notwithstanding that, however, the approach does seem to make some counter-intuitive predictions (see Carston 2002*b*). To maintain his distinction

[14] Sperber 1995, p. 195 writes 'when most of us talk of reasoning, we think of an occasional, conscious, difficult and rather slow activity. What modern psychology has shown is that something like reasoning goes on all the time—unconsciously, painlessly and fast.' It is in this latter, broader sense that relevance theorists regard linguistic communication as *inferential*. As I understand it, Recanati's claim is that the relevance theory notion of *broad* inference actually falls somewhere *between* what Recanati takes to be more satisfactory definitions of inference in the broad and the narrow sense. He goes on to argue that primary pragmatic processes may only be regarded as inferential in a *very* broad sense—in the manner we might regard processes underlying perception (e.g. vision) as inferential—a sense in which, he argues, they cannot *properly* be regarded as inferential.

between primary and *secondary* pragmatic processes (i.e. those processes responsible for the interpretation of cases in which a speaker means something over and above what they say) Recanati utilises a notion originally due to Millikan (1984). According to Millikan, in those cases in which there is some divergence between what a speaker says and what a speaker means, communicators are effectively 'tinkering' with the 'mechanisms of normal language flow' (1984, p. 69). By 'normal' language flow Recanati is referring to what a speaker *says*, in the sense of the truth-conditional proposition expressed by the speaker. But at what point (and by what criteria) do we decide that language flow is *not* normal? As Carston shows, communicative exchanges involving cases of implicature often appear more 'normal' than those in which the implicature is fully spelt out; the latter often appear clumsily put, or inappropriately over-explicit.

More importantly, given the concerns in this thesis, we have seen in Chapter Three that the interpretation of patently *non*-linguistic, clearly *non*-coded natural behaviours—shivers, for example—plays a crucial role in an audience's derivation of what a speaker has *said*. On any conception of 'normal', we would presumably want to say that the mechanisms underlying their interpretation are part of the mechanisms of normal language flow; but if they are not coded, then it is hard to see how (on any conception of 'inference') these interpretive mechanisms are not inferential.

Conspicuously absent from my discussion of the chase between A and B above is the question of how the facial expressions (and other *coded* natural behaviours) of the two individuals would function as a window into their respective emotional states. It is clear that they would also play a large part in

your interpretation of the situation; on the basis of their appearance, you would sense not only A and B's thoughts, but also their feelings; that B is (more or less) frightened, and A is angry (or furious).

Though it is plainly involved to some extent, the degree to which the wider mind-reading ability is implicated in the reading of such emotional states is unclear. If, as I am proposing, some of the behaviours that indicate emotional states are coded behaviours, then we might predict that they are interpreted via their own dedicated cognitive mechanisms. What *is* clear, though, is that the reading of the faces of others is sure to be problematic to those for whom the mind *behind* the face is out of reach.

Here too, research into autism is suggestive. There is a growing literature charting the difficulties that autistic subjects have in attributing emotional states to others (Hobson, Ouston and Lee 1988; Muris, Meesters, Merckelbach and Lomme 1995). In an attempt to shed more light on the issue, Baron-Cohen, Spitz and Cross (1993) examined the recognition of emotion in autistic children. Based on the observation that autists have problems in recognizing beliefs, Baron-Cohen *et al* theorized about the extent to which this would manifest itself in the recognition of 'cognitive' emotions. These are emotions such as surprise which, since they are caused by beliefs, presume some sort of understanding of beliefs. Baron-Cohen *et al* regards these as distinct from 'simple' emotions—those emotions caused by situations (such as happiness and sadness). As predicted, the autistic children had more difficulty recognising surprise.[15]

[15] It's not entirely clear to me whether this distinction holds, since just as I can by 'happy', I can also be 'happy that *P*'. The point, I suppose, is that there is a sense in which you can only be 'surprised that *P*'; that is to say, you can't just feel 'surprised'. It might also be argued, however,

To the extent that these findings (and the suppositions on which they are based) are correct, it might be taken to suggest that while the meta-communicative and meta-psychological abilities in these subjects are impaired, certain of the natural coding-decoding mechanisms remain intact. This may pave the way to shedding light on the precise nature of the relationship between the wider meta-psychological mindreading ability and those abilities responsible for the coding-decoding of behaviours linked to natural codes. As we saw in the previous chapter, humans do have neural mechanisms dedicated to processing facial expressions (Gazzaniga and Smiley 1991). These specialised mechanisms, like those meta-communicative abilities that govern the search for relevance in verbal comprehension, may also be sub-modules of the wider theory of mind ability.

One issue that has come out of my research on facial expressions is that the notion of intentionality (in the 'rich, philosophical sense'—see *fn.* 2 from my introduction) is largely (and conspicuously) absent from the literature. Fridlund (1994) abstracts away from it entirely: 'I have circumvented these "levels of intentionality" issues in the interests of space, and use intentionality in a purely functionalist sense.' (p. 146)

Of course, we underestimate the power of human facial expression at our peril too. It is certainly true that '...whereas the prosimian face is relatively unexpressive, the monkeys and apes tend to exhibit a quite significant range of expressions, culminating in the Marcel Marceau of expressiveness, modern humans' (Hauser 1996, p. 265). But if there is one moral to draw from my research, it is that what distinguishes human from non-human animal

---

that 'emotion' is not the correct word to use for simply 'feeling' happy (see earlier discussion of Rey's distinction between 'feelings' and 'emotions' in Chapter Four, p. 159).

communication (and interaction generally) is not that we live in a world populated by other faces, but that we live in a world populated by other minds.

2. *Two* 'showing-meaning' continua

In Chapter Four I discussed Erving Goffman's 'response cry' approach to interjections. Interjections, Goffman writes, 'can't quite be called part of language' (1981, p. 115); however, they exist on a continuum between display and language proper: 'response cries such as *eek!* might be seen as peripheral to the linguist's domain ... but imprecations ... are more germane, passing beyond semiword segregates to the traditional material of linguistic analysis' (1981, p. 121).

Goffman is not the only other researcher for whom the notion of some sort of continuum between display and language proper, or showing and saying, is intuitively appealing. David McNeill (1992) presents another such continuum, based on one originally presented by Adam Kendon (1988). Although this continuum focuses on gestural rather than vocal behaviours, there are clear parallels to be drawn between it, Goffman's continuum and the continuum that has emerged as a thread running through this thesis. Kendon's idea is that as we move along the continuum, the behaviours take on more 'language-like' properties, and 'Kendon's continuum'[16] is reproduced below (from McNeill 1992, p. 37) in *fig. 4*:

[16] So named in Kendon's honour by McNeill (1992, p. 37).

fig. 4

Kendon's Continuum

Gesticulation → Language-like gestures → Pantomimes → Emblems → Sign Languages

Those movements classified as 'gesticulation' in the continuum are those movements of the arms and hands that spontaneously accompany speech; so when Jack utters 'the dentist had a good poke around' he might gesture with his forefinger vaguely in the direction of his open mouth. 'Language-like' gestures are similar to gesticulations but are, in one respect at least, 'integrated' into a linguistic string; so Jack might utter 'the examination was OK, but when he started [*gesture to represent drilling*] it was agony'.[17] 'Pantomimes' are those movements that depict objects or actions; accompanying speech is no longer obligatory—'there may be either silence or just inarticulate onomatopoetic [*sic*] sound effects ("whoops!", "click!" etc.).' Again parallels with Goffman's analysis of interjections are germane. 'Emblems'[18] are gestures that McNeill refers to as 'Italianate': the British two-fingered insult would be a case (though, presumably, a *non*-Italianate one). Finally, Sign Languages are, of course, languages proper, with their own syntactic, semantic and phonological[19] rules.

[17] I'm not convinced that 'Language-like' is the clearest terminology to use for this type of behaviour.

[18] McNeill (p. 56) credits Ekman and Friesen (1969) with an earlier use of the term 'emblem'; Ekman and Friesen (1969) credit Efron (1941) with the original.

[19] I'm assured by Neil Smith (p.c.) that sign languages *do* have their own phonological rules—so instead of distinctive features, there is a set of parameters such as hand-shape, location, movement, orientation of the palms relative to the body, and facial action. The idea is that these function in much the same way as the auditory parameters of spoken language, with minimal pairs etc.

In Chapter Four I was concerned with attempting to draw parallels between the 'showing-meaning' continuum as presented in Chapter Three, and the other continua mentioned at various points and outlined above. However, as I also remarked, there are differences. The aim of the following discussion is to focus on these; for while the two continua reflect similar underlying intuitions, in key regards they are actually quite distinct. My conclusion will be that while continua of the kind proposed by Kendon and Goffman are useful descriptive tools, and capture the intuition that a continuum of some sort does exist between display and language proper, they are lacking in any real explanatory power. By contrast, the showing-meaning$_{NN}$ continuum developed in this thesis is useful in both descriptive and explanatory terms.

Continua of the kind proposed by Goffman and Kendon are based entirely on the role played in communication by *coding*:[20] in semiotic terms, the nature of the 'sign'. The principal observation on which the continuum is based is that as we move across the full spectrum of behaviours from display to language, there is an increase in 'language-like' codification. Among the features typical of such 'language-like' coding, McNeill lists 'segmentation', 'compositionality', a 'lexicon', a 'syntax', 'distinctiveness' and 'arbitrariness'; these features are highly reminiscent of Hockett's list of the distinguishing features of language discussed briefly in the previous chapter.

It should be noted, however, that although the point of such a continuum is that there is an increase in 'language-like' coding as we move from left to right, the ultimate aim appears to be to analyse *most* (if not all) the behaviours along the continuum as *codes* of one sort or another:

[20] Allowing for my earlier point that Goffman is not explicit as to whether his continuum is a 'coding' continuum.

> 'The first point to establish is that, in performing gestures,[21] the speaker's hands are no longer just hands, but *symbols*. Gestures are not just movements and can never be fully explained in purely kinesic terms. They are not just the arms waving in the air, but *symbols that exhibit meanings* in their own right.'
>
> (1992, p. 105)

> 'Our purpose is thus to bring out semiotic values, and this has led us to build semiotic distinctions directly into the gesture classification; that is, to classify the gesture by means of asking (a) is the movement a symbol? and (b) what type of symbol is it? The categories of iconic, metaphoric and deictic, or beat correspond to the fundamental types of semiotic sign.'
>
> (*ibid.* p. 77)

For obvious reasons, I call a continuum such as this—which clearly has its roots in the code model of communication—a *C*ode-continuum, or a *C*-continuum. Essentially, a *C*-continuum is a continuum between non-linguistic and linguistic coding.

By contrast, the continuum proposed in Chapter Three is based on the role played in ostensive communication by the inferential attribution of intentions; it has its roots in Gricean notions such as those discussed in Chapters One and Three. Positions on the continuum reflect the extent to which hearers are required to attribute intentions to speakers[22] in order to get from the evidence provided to the information being communicated (the first, basic layer). I call this continuum an *O*stensive behaviour-continuum, or an *O*-continuum.

---

[21] By which McNeill means 'gesticulations' on Kendon's continuum; 'I use the term "gesture" in this book specifically to refer to the leftmost, "gesticulation" end of the spectrum.' (1992, p. 37).

[22] Or producers of ostensive acts generally.

My claim is that the *O*-continuum can be extended in such a way as to incorporate the *C*-continuum; after all, it follows from there being a high degree of coding in the communicative stimuli that the evidence provided will be of the less direct variety.[23] The *C*-continuum, however, cannot be extended to incorporate the *O*-continuum.

One thing that the *C*-continuum and the *O*-continuum have in common is that they provide a 'snapshot' of the types of evidence used in communicative acts. So at one extreme of both continua lie clear cases of spontaneous, natural display (though in the case of the C-continuum, these are also coded signals); while at the other extreme lie clear cases of linguistic coding. In between lie a range of cases. However, while both continua illustrate the increase in 'language-like' codification as we move toward language proper, there are a variety of points about which the *C*-continuum can apparently say little (if anything). This, I argue, is because it takes no account of the inferential nature of human communication.

Because the *C*-continuum is based on the code model, and concerned with 'classifying signs', it can say little about *how* communicative behaviours are used

---

[23] Some people have objected to my characterisation of a linguistic utterance as only providing *in*direct evidence of the first, basic layer of information—largely, I suppose, on the strength of the observation that as native speakers of a given language we do indeed feel that linguistic utterances are 'direct' evidence, in the sense that we understand them immediately and effortlessly. This harks back to the earlier discussion of the Recanati/Millikan view of linguistic communication as a form of perception (see pp. 221-223 above). The point, I suppose, is that they only feel direct *if you know the code*. If, for example, a speaker of Burmese were to utter a sentence of Burmese to me, I would not feel that the evidence she was providing was in the least bit direct, since it would provide me with no evidence of the first, basic layer at all (other than the fact that she had obviously said something). By contrast, if the Burmese person in question were to smile at me, I would understand perfectly well: that signal forms part of what (I presume) is a code we all know—a *universal* natural code.

other than to say that they are coded signals. The positions of various behaviours on the *C*-continuum are fixed or static: facial expressions 'display'; words 'mean'; interjections belong at some half-way point. On the face of it, this seems fair enough, but a moment's reflection reveals that it captures neither the way people communicate, nor the manner in which behaviours are used.

For example, it misses the crucial point that just as communicators use language to *say* things, we can also use language to *display* or *show*. Equally, communicators can use what are essentially 'displays' to $mean_{NN}$. In the *O*-continuum there is complete freedom for stimuli to move along the continuum depending on how they are used.

Consider a case in which I intend to communicate to you that I have a sore throat. I might choose to communicate this to you by saying (1) in an extremely hoarse voice:

(1) The moon looks beautiful tonight.

Or another case, in which Lily approaches Jack at a party and asks him if he speaks any French; Jack replies as in (2):

(2) Pouvez-vous parler moins vite s'il vous plaît?

Both these qualify as clear cases of showing or display, since direct evidence of the first, layer of information is provided—in (1) that I have a sore throat, in (2) that Jack can indeed speak French (see Sperber and Wilson 1986/1995, pp. 177-178).

Consider ostensive uses of natural signs, such as those discussed in Chapter One: for example, the case in which Lily uses a shiver to $mean_{NN}$ that she does not want to stay outside. Here, as I suggest above, a behaviour that is essentially a display is used (in virtue of the intentions behind the display) to $mean_{NN}$. Though the *O*-continuum also provides a snapshot, that is not all it tells us; it suggests not only how the behaviours that fall along it are interpreted (via a mixture of direct and indirect evidence—or, if you prefer—natural and coded behaviour), but how it is that they might, at various times, occupy different points.

In Chapter Five I suggested that since we have a continuum, we would expect expressions to move along it. In historical terms, when an interjection, for example, moves far enough along the continuum it may become linguistically productive ('to wow', 'yucky'), and some of its uses may be properly linguistic (verbs, adjectives etc.). This suggests a historic, diachronic dimension to the continuum. A *C*-continuum, however, can shed little light on this dimension. Since behaviours occupy fixed positions on the continuum, it can only, at best, represent diachrony in a series of still pictures.

By contrast, the *O*-continuum, which can *also* been seen as a series of still pictures, has the added advantage that it can represent the fluidity and constant change that results in expressions coming to form part of language. In many accounts of historical linguistics (Aitchison 1991, Lightfoot 1991), children converge on the simplest grammar that reflects the practice of the speech community to which they are exposed. The continuum may allow us to explore the idea that pragmatic factors may affect this convergence, to see language change in terms of the *micro*-processes involved in the emergence of new

encoded meanings. Language change might then be characterised as the population-scale *macro*-processes that are the result of the accumulation of those micro-processes, and hence result in the stabilization of new senses.[24]

Presumably, there is also some degree of interaction between cognitive and social factors, which influence the direction of linguistic change. Dan Sperber (1996) uses the notion of 'attractor'[25] in what he terms 'epidemiological' models of ephemeral and longer-lasting *cultural* change. Whilst this notion has been discussed in the context of historical linguistics (Lass 1997), an epidemiological model has not yet been applied to historical linguistics and could shed useful light on the distinction between short-term linguistic fashions and longer-lasting trends, and the causal processes that influence them.

As well as the historic-diachronic dimension to the continuum, we might also explore whether it has an *evolutionary*-diachronic dimension. Crucially, when it comes to evolutionary concerns, the *O*-continuum is *doubly* suggestive. Firstly, as we have seen, the flexibility inherent in the continuum allows for the fact that behaviours might move along it; secondly, and perhaps more importantly, by using the *O*-continuum to shed light on the evolution of language and communication, we are setting our account within the wider context of the evolution of human cognitive abilities. I turn to this in the next section.

---

[24] Which—following a suggestion by J. L. Speranza (p.c.)—I would like to call *Quantum Pragmatics*.

[25] 'To say that something is an attractor is just to say that, in a given space of possibilities, transformation probabilities for a certain pattern: they tend to be biased so as to favour the transformations in the direction of some certain point, and therefore cluster at or around that point. An attractor is not a material thing; it does not physically "attract" anything. To say that there is an attractor is not to give a causal explanation; it is to put in a certain light what is to be causally explained' (Sperber 1996, p. 112). Various factors can be responsible for attractors—cognitive, social, environmental—each operating on different time-scales.

3. A prince among primates

> '...language must have begun from attempts at communication between a few individuals. At first these efforts at communication did not have very much stability of literal meaning. Only slowly and after much time did a stable community of users lead to the abstract concept of literal meaning ... There is no hard and fast platonic literal meaning that utterers' meanings attach themselves to ... The story surely is exactly the other way round.'
>
> (Patrick Suppes 1986, p. 113)

Over recent years, there has been a tremendous resurgence of interest in the evolution of communication, cognition and language. In three fairly recent books on the evolution of language, three authors each regard a different aspect of human language as the evolutionary cornerstone around which the edifice of the language capacity might have been built. For Terrence Deacon (1997) it is the fact that language uses 'symbols'; for Jean Aitchison (1996) it is what she refers to as 'the naming insight', reflected in the early stages of the acquisition of language by children; for Andrew Carstairs-McCarthy (1999) it is subject-predicate structure, which he argues has its origins in syllable structure.[26]

What is interesting to note is that all these authors choose to base their account on a view of ancestral hominid linguistic communication which was a simple coding-decoding affair. The same stance is adopted in what is regarded as one of the seminal papers on the topic: Steven Pinker and Paul Bloom's 'Natural language and natural selection' (1990). Such a view, however, is problematic

[26] Bickerton (2000) takes issue with Carstairs-McCarthy's proposal, claiming that since syllable structure lacks recursion, the tripartite structure of the syllable—[(ONSET) [NUCLEUS (CODA)]]—can not have been a precursor to syntax, or, more precisely, [(SUBJECT) [VERB (OBJECT)]] structure. This may indeed be true, but there is a far more worrying problem with Carstairs-McCarthy's proposal. That is, that there is a wealth of convincing evidence suggesting that, actually, syllables do not have [(ONSET) [NUCLEUS (CODA)]] structure at all, and that what have previously been called 'codas' actually exist in the 'onset' position of a following syllable containing a null vowel (see Harris and Gussman 2002).

(see Sperber 2000): it is hard to see why, if it began as a pure coding-decoding process, at a certain evolutionary stage human linguistic communication should have changed in character so drastically.

Another tacit assumption in the works cited above is that the human linguistic ability preceded the human metarepresentational ability.[27] This view is also problematic: firstly, it is hard to see how humans could ever have become aware of the representational character of their signals; secondly, and perhaps more importantly, under such a view it is hard to see how a language faculty could have been adaptive.

For one of the key problems in formulating a plausible evolutionary account of the language faculty based on natural selection is that a mutant on whom some novel linguistic coding and decoding ability was conferred, and who was in a community which had no capacity for inferential communication, would have had no advantage in 'fitness'. Her abilities would have been entirely useless; she would, after all, have had no one to talk to, and no one to listen to. However, the problem is more tractable if we hypothesise that metarepresentational abilities *preceded* linguistic abilities. Humans with such abilities would have been involved in inferential communication before the evolution of a language faculty as we know it, and a plausible evolutionary scenario presents itself: the biological evolution of the language faculty can be regarded as involving the evolution of ever more precise coded signals, which make the input to *existing* inferential processes involved in existing human communication more efficient. If humans were communicating inferentially before the development of language, then showing—in the sense described in Chapter Three—clearly came

[27] Deacon is actually quite specific on this matter; he proposes that the human capacity for symbolic thought could have led to Theory of Mind.

before saying, and evolutionary considerations might provide us with useful insights into why modern humans are so adept at doing both, and why there should exist a continuum between the two.

Notably, the metarepresentation-before-language hypothesis is independently endorsed by evolutionary psychologists, primatologists and ethologists working on the evolution of human social intelligence, who argue that the social pressures of ancestral hominid life would have led to the evolution of metarepresentational abilities either to outmanoeuvre opponents, or detect cheats in the community (Cosmides 1989).[28] The need to manipulate others may have been one of the main factors in the evolution of communication, and the recent literature on *Machiavellian Intelligence*—Byrne and Whiten 1988, Whiten and Byrne 1997—is particularly instructive in this regard. In the words of Nicholas Humphrey:

> 'Once a society has reached a certain level of complexity, then new internal pressures must arise which act to increase its complexity still further... If intellectual prowess is correlated with social success, and if social success means high biological fitness, then a heritable trait which increases the ability of an individual to outwit his fellows will soon spread through the gene pool. In these circumstances there can be no going back; an evolutionary "ratchet" has been set up, acting like a self-winding watch to increase the general intellectual standing of the species.'
>
> (Humphrey 1988, p. 21)

[28] To complement this, there is also an increasing body of work which attempts to corroborate the claim that *some* kind of mind-reading capacity was indeed a necessary cognitive precursor to language (see, for example, Dunbar 1998).

Interestingly, this kind of view does not fit with the traditional philosophical view of the function of communication. Recall from Chapter One that Neale's main concern over Grice's modification of his 1957 characterisation of meaning$_{NN}$—which, recall, led him to distinguish between *exhibitive* and *protreptic* utterances—was that 'it does not comport well with the commonly held view that the primary purpose of communication is the transfer of information about the world' (1992, p. 549); on Grice's revised account, Neale goes on, 'the primary purpose seems to be the transfer of information about one's mental states' (*ibid.*).

Aspects of a recent paper by Dan Sperber (2001) might be seen as militating against this view. In the case of honeybees, the 'primary purpose' of communication does indeed appear to be to transfer information about the world. Having perceived a source of nectar, the forager bee returns to the hive and performs its dance. Once the bee's dance has been decoded, the receiver bees are presumably in precisely the same cognitive (mental) state as the honeybee that has performed the dance was when it discovered the source. Communication is beneficial to the bees insofar as they benefit from the perceptions of other bees; it results in what Sperber calls 'cognition by proxy'.

As we saw in Chapter Three, human inferential communication does not result in the straightforward replication of cognitive states that we see in bees. As Hauser (1996, p. 497) puts it: 'whereas bees are informational laser beams, humans are informational floodlights'. Sperber points out that even when she *is* trying to transfer information a human communicator is not always benevolent. If it is in her interests, and will help her to achieve her goals, she is capable of lying and deception. While there is some evidence to suggest that some non-human

primates are capable of deception in a basic, stereotyped manner, humans are capable of highly elaborate, novel deceptions. Equally, of course, audiences are not always trusting: a communicator's motives may be entirely benevolent—indeed, she may be wholly honest—but she may still not be believed.[29]

A plausible account of the function of communication must take account of these observations; furthermore, a plausible account of the evolution of communication must take account of the fact that despite this, and despite the fact that a communicator and her audience may have conflicting goals, communication must have been advantageous to *both* in order to stabilise (or be selected for).

Sperber argues that the human ability to present or assess coherent argumentation may have been one of the factors involved in this stabilisation: for communicators, argumentation is a means of persuasion; for an audience, it is a means by which they can assess the content of a message, irrespective of the degree of trust they have in the person who is communicating with them. Indeed, Sperber suggests that the evolution of the human capacity for *reason* may well have met communicative ends, rather than those of individual knowledge acquisition. Communicated information already comes 'tagged' with higher-level mental states (or propositional attitudes), within which the basic information

[29] Actually, there is some evidence that honeybees are able—to some degree at least—to maintain a level of scepticism about the dance of a returning forager. In an ingenious experiment, Gould (1990) removed some forager bees from a hive and introduced them to a source of pollen on a land-bound boat. He then prevented the bees returning to the hive. Over time, he moved the boat further and further into the middle of a nearby lake (forcing the bees to fly further and further) and only allowed them to return to the hive and perform their dance when the boat—and therefore the pollen source—was situated right in the middle of the lake. When the foragers did indeed return to their hive and dance, *none* of the bees in the hive visited the boat. Gould's interpretation of this was that the bees were comparing the supposed location of the pollen with their knowledge of their environment, and (in a manner of speaking) not *believing* the dancers.

being transferred is embedded: a human communicator communicates that she believes, desires or regrets that *p*; an audience responds that that he doubts or fears that *p*. Checking the coherence of argumentation would have required an even more sophisticated metarepresentational ability.

Whatever the implications for the evolution of the human reasoning ability, notice that construed in this way, human communication is at least as much about the transfer of information about mental states, as it is about states of affairs *per se*. This added dimension in human communication may have been part of Humphrey's evolutionary 'ratchet' that helped the spark the cognitive arms race[30] from which natural language finally emerged.[31] Language would have been a perfect adaptation to increase the efficiency of inferential communication: the vehicle by which thoughts and ideas are carried, and through which they proliferate.[32]

In many ways, the most surprising contribution to the debate on the evolution of language is that of Hauser, Chomsky and Fitch (2002).[33] Hauser *et al* seek to sever the link between the evolution of language and the evolution of

---

[30] In honeybees, of course, this cognitive arms race never took place. This may be one reason why honeybees are pretty much exactly the same as they were eighty (or two hundred and twenty—see Demko 1995) million years ago, and humans have evolved into the sophisticated, farsighted creatures we know and love today.

[31] It is not Sperber's aim in his 2001 paper to present an account of the evolution of language, but he does also suggest that as argumentation evolves, so a logical vocabulary—words such as 'if', 'however', 'so'—would be beneficial; these expressions may have been the evolutionary forerunners of the *non-translational* vocabulary.

[32] If, in humans at least, communication is causal in this way, then it raises another question. In the case of human natural codes, do we—for example—smile to let others know that we are happy (i.e. are they, in Grice's terms, *exhibitive*), or do we smile to make others happy too (i.e. are they *protreptic*)?

[33] Which may be one of the reasons why, at UCL at least, the 'Evolution of Language' course is run by a non-linguist within the Department of Anthropology.

communication. They argue that language—in the sense of a 'narrow syntax' (2002, p. 2) which generates linguistic representations and maps them on to the conceptual-intentional and sensory-motor interfaces—could have evolved as a by-product of other human computational abilities: in evolutionary terms, it may be a 'spandrel' (Gould and Lewontin 1979).[34] One of the motivations for severing the link, and the main factor behind their belief that 'investigations of this capacity should include domains other than communication (e.g. number, social relationships, navigation)' (2002, p. 2) is that 'the core recursive aspect of FLN [faculty of language in a narrow sense, TW] appears to lack any significant analog in animal communication' (*ibid.*). By contrast, features of what Hauser *et al* call 'FLB' (faculty of language in the *broad* sense—*including* the conceptual-intentional and sensory-motor systems) do appear to have homologues in non-human animals.

These arguments, however, do not militate against the metarepresentation-before-language view presented above. Firstly, the proposal that humans were capable of entertaining representations such as, for example, 'She meant P' or 'She intended me to believe P' before the evolution of FLN is not inconsistent with FLN having emerged entirely as a spandrel. Secondly, and more

[34] In Gould and Lewontin's words, the *spandrels* of San Marco Cathedral are 'the tapering triangular spaces formed by the intersection of two rounded angles at right angles' and 'are necessary architectural by-products of mounting a dome on rounded arches'. Each spandrel 'contains a design... so elaborate, harmonious and purposeful that we are tempted to view it as the starting point of any analysis.' However, 'this would invert the proper path of analysis', since 'the system begins with an architectural constraint: the necessary four spandrels and their tapering triangular form'. Gould and Lewontin go on, 'anyone who tried to argue that structure [spandrels] exists because of [the designs] would be inviting the same ridicule that Voltaire heaped on Dr. Pangloss... Yet evolutionary biologists, in their tendency to focus on immediate adaptation to local conditions, do tend to ignore architectural constraints and perform just such an inversion of explanation' (Gould and Lewontin 1979, pp. 147-149).

importantly, nor is the claim that FLN evolved partially as a by-product of other abilities necessarily inconsistent with its having evolved to meet largely communicative ends. For while it is certainly true that 'the core recursive aspect of FLN appears to lack any significant analog in animal communication' it is not true that it is the only *human* recursive ability. The human metarepresentational ability is a recursive, syntactic ability *par excellence*; a plausible candidate, it might be argued, for exaptation into syntax in Hauser *et al*'s narrow linguistic sense. The level of metarepresentation required for inferential communication presupposes a considerable recursive ability, and the syntax of FLN could quite plausibly have been inherited from the syntax of the language of thought. As Tooby and Cosmides point out in their commentary to Pinker and Bloom's 1990 article, just as we should not ignore architectural constraints, we should be wary of 'naïve spandrelism'.[35]

Precisely what kind of existing behaviours might have been co-opted into the service of inferential communication, and hence laid the foundations for the emergence of public language, is unclear. There are a number of possibilities. An existing repertoire of coded vocal signals, such as the warning calls of modern day vervet monkeys is one possibility; another is that in tandem with vocal calls, ancestral communication may well have involved the use of gesture, facial-signalling or mime (Donald 1998). Yet another possibility is that the source is instinctive emotional calls. This last option is one endorsed by Ray Jackendoff, who suggests that interjections—those semi-words discussed in Chapter Four—might represent '*fossils* of the one word stage of language evolution' (1999, p.

[35] Origgi and Sperber 2000 devote the last few pages of their paper to an account of how a rudimentary syntax might have evolved.

273).[36] Indeed, there may be evidence that interjections do represent some kind of more primitive communicative system:[37] in neurological terms, use of interjections is associated with the phylogenetically ancient limbic sub-cortical circuitry linked with emotion, as opposed to the more recent cortical structures implicated in the production of language proper.

Each of these proposals is, however, fraught with problems, the main one being that there is, of course, none of the evidence that is usually used to confirm (or disconfirm) evolutionary hypotheses: e.g. a fossil record. It may be that, for the moment at least, we will have to content ourselves with a myth.

---

[36] Proposals of this kind have not met with universal appeal: '[T]he universals of language are so different from anything else in nature... that origin as a side consequence of the brain's enhanced capacity, rather than as a simple advance in continuity from ancestral grunts... seems indicated (Stephen Jay Gould 1989, p. 14). In this view Gould has an unlikely ally—given his own views on 'language'—in Chomsky himself: 'in the case of such systems as language or wings it is not easy even to imagine a course of selection that might have given rise to them' (1988, p. 167).

Dennett (1996) discusses Gould's and Chomsky's scepticism over the role natural selection may have played in the evolution of language:

> 'Gould and Chomsky...float the suggestion that nothing we know yet rules out the possibility that... change in brain size... could have as an adventitious consequence radical discontinuities in behavioural repertoire (hint: such as the sudden blossoming of a Language Acquisition Device). Right. And nothing we know yet rules out the hypothesis that given a few lucky mutations and a slight change in their diet, pigs may suddenly sprout wings or start spinning magnificent pigwebs for the first time in their biological history' (Dennett 1996, p. 264).

[37] *Cf.* Jesperson (1922, p. 414) '[One] theory is the interjectional, nicknamed the *pooh-pooh*, theory: 'language is derived from instinctive ejaculations called forth by pain or other intense feelings or sensations'.

4. Myths

'The three of them stood and looked at each other. Then, as so often happened with the people, there were feelings between them. Fa and Nil shared a picture of Ha thinking.'

(William Golding—*The Inheritors*)

Grice (1982) presents us with just such a 'myth' of how human cognitive capacities might have spiralled in such a way that $meaning_{NN}$ might have emerged from $meaning_{N}$. He goes on:

> 'But can such a link be explained by a *myth*? The question is perhaps paralleled, as was recently suggested to me, by the question how the nature and validity of political obligation (or perhaps even of moral obligation) can possibly be explained by a mythical social contract.'
>
> (1989, p. 297)

Would that Grice's scrupulousness were observed by all researchers looking into evolutionary issues. Much work on the evolution of communication and language begins with the assumption that since there presumably is an evolutionary explanation, any evolutionary explanation is probably worth pursuing. Moreover, as I mentioned above, there is a strong case for suggesting that in the case of brains (which do not fossilize), and languages (which, unlike linguists, do not fossilize either) a myth is the best we have to offer at the present time.

Grice asks us to imagine a creature;[38] a creature which, when in pain, involuntarily emits a noise—for the sake of argument, a groan.[39] At this stage,

[38] Much of Grice's work in philosophical psychology involves—what he calls—'creature construction'. That is, we imagine ourselves as benevolent *genitors*, designing and constructing organisms (or *operants*) so that their chances of survival might be maximized (Grice is not solely

we might say—in a manner entirely analogous to those examples in Chapter One—that the groan means$_N$ that the creature is in pain. Grice then moves through a succession of four further stages designed to show what needs to be added to this scenario in order to have what we deem to be a case of meaning$_{NN}$.

In the second stage, Grice notes, we require that firstly, creature X is able to produce the behaviour in question voluntarily and secondly, that another creature—creature Y—recognizes the voluntary nature of the sound X has produced. At this stage, of course, the fact that Y recognizes the fact that X has produced the piece of behaviour voluntarily will most likely lead Y to 'interpret' X as *not* being in pain (since X is producing a normally involuntary noise voluntarily).

At the next stage, we imagine that not only does Y recognize that X has produced the noise voluntarily, but also that X *intends* Y to recognize his behaviour as such. The possibility of deception—a plausible interpretation in the previous stage—is no longer the only possibility:

> '...we have now undermined the idea that this is a straightforward piece of deception. Deception consists in trying to get a creature to accept certain things as signs of something or other without knowing that this is a faked case. Here, however, we would have a sort of perverse faked case, in which something is faked but at the same time a clear indication is put in that the faking has been done.'
>
> (1989, p. 293)

---

concerned with survival *per se*, but for present purposes that will do). I concentrate on the small section from Grice (1981) rather than the longer, considerably more dense published work from 1975 (published in 2000) for reasons of space (my own psychic-conceptual space, that is).

[39] The groan could just as easily be a spontaneous gesture, or even a natural coded behaviour, such as an alarm call.

At this stage, of course, Y is likely to be in something of a quandary. Creature X is not only simulating that he is in pain, but simulating that he is in pain overtly. It does not seem unreasonable to suggest that one way Y might proceed is to presume that X is 'engaging in some form of play or make-believe' (*ibid.*).[40] By the time such a conclusion is arrived at by Y, we have reached stage four.

By the fifth stage, Y has come to suppose not that X is playing a game, but rather that X is trying to get Y to believe (or at least accept) that X is in pain; that is, that X's intention in producing the normally *in*voluntary behaviour voluntarily is to communicate the very same thing for which the *in*voluntary behaviour is a natural sign. The idea may seem far-fetched, but of course it is not. As just one example, consider how often fake yawns are used to communicate tiredness. By this stage, of course, Y may wonder precisely why X should choose to use a faked expression of pain to communicate pain as opposed to an involuntary sound. Grice suggests various reasons; that it might be 'uncreaturely' to act so spontaneously, or (more likely) that X's voluntary production might only be intended to indicate some, as opposed to all, of the features associated with spontaneous emissions. To this we might add the possibility that, following the discussion in Section 3 of Chapter One, X may have good reason to 'show' an *in*voluntary behaviour. In this regard, consider, the earlier examples in which Lily shows Jack her spontaneous smile, or her spontaneous shiver. Grice, of course, does not discuss this possibility, since he was interested in characterising meaning_{NN}, which for him could not involve the deliberate showing of

[40] The myth is only designed to shed light on phylogeny; nonetheless, in the light of Leslie's proposal that pretence is the first epistemic mental state that children come to recognise and use (if pretence can be regarded as an epistemic mental state), it may be that it has ontogenetic implications too.

spontaneous natural behaviours (see earlier quote) or, indeed, anything that provided direct evidence of its own for a certain state of affairs.[41]

By the time we have reached the fifth stage, the intention behind the behaviour, rather than the behaviour itself, plays a central role in Y's successful understanding of X, and we have now reached:

> 'a stage in which the communication vehicles do not have to be, initially, natural signs of that which they are used to communicate; provided a bit of behaviour could be expected to be seen by the receiving creature as having a discernible connection with a particular piece of information, then that bit of behaviour will be usable by the transmitting creature, provided that the creature can place a fair bet on the connection being made by the receiving creature.'
>
> (1989, p. 296)

Once we reach the fifth stage of Grice's explanation, the vehicles of communication (and communication itself) can be characterised in terms of the *O*-continuum. We have reached full-fledged ostensive communication as characterised in earlier chapters. Were we tempted, we might extend Grice's myth and imagine further scenarios in which the evidence provided by a communicator becomes less and less and direct, with successful communication depending to an ever greater extent on attributing intentions to X. The stimulus itself may begin to consist of *ever more* stylised imitations of spontaneous

[41] As we saw earlier, and in contrast to Grice, Schiffer (1972) was happy to describe some of those cases of 'deliberately and openly showing' as meaning_{NN} (the case of the bandaged leg, for example). It's unclear whether he would have made the same of intentionally shown spontaneous behaviours.

emissions, and might thus be seen as occupying various points along the continuum.

At a given point, the success of communication need no longer depend on any prior *natural* connection between the ostensive stimulus and the intended meaning, but perhaps instead on some prior *stylised* version of that connection. What I have in mind is something analogous to the development of writing systems, in which representations that were originally iconic become increasingly stylised. Over the course of time, the final figure (or letter) that emerges as a result of historical processes bears no resemblance at all to the *object* that was originally represented by the stylised representation, but continues to bear some relation to the original representation itself. In such a way the increase in codification discussed in Chapter Four, and illustrated by the *C*-continuum, is accounted for in terms of increased reliance by the audience on the intentions behind a stimulus. The *C*-continuum is thus accommodated within the *O*-continuum.

Of course, this *is* only a myth; I am no *pooh-pooh theory* revivalist. It is difficult—not to say, impossible—to imagine how full-blown language might have emerged from an accumulation of basic communicative exchanges, how something so complex and sophisticated could ever have crystallised from such beginnings. Here, Hauser *et al*'s 'spandrel' account may be the best we have, and we may well have to accept that we will never know the true nature of the forces which shaped the emergence of all the components of language proper.

As Sperber and Wilson put it:

> '[Grice's myth] is reminiscent of the story of how Rockefeller became a millionaire. One day, when he was very young and poor,

> Rockefeller found a one-cent coin in the street. He bought an apple, polished it, sold it for two cents, bought two apples, polished them, sold them for four cents... After one month he bought a cart, after two years he was about to buy a grocery store, when he inherited the fortune of his millionaire uncle.'
>
> (1986/1995, p. 53)

Nonetheless, if communication did play a role in the evolution of public language, the myth is still illuminating. Firstly, as I mentioned earlier, those metarepresentational—in a sense, *syntactic*—abilities that underlie inferential communication are just the kind of abilities that may have been exapted for the syntax of natural language. Secondly and more generally, it doesn't matter what kind of behaviours were used in early inferential communication—existing coded warning calls, instinctive emotional calls, mimes, gestures, facial expressions; what we have is a plausible, naturalistic framework around which to build.

It is hard to see how insights such as this are possible within semiotic, code-based accounts. As I mentioned in Chapter Five, the semiotic view does not sit comfortably with the Chomskyan view of the human ability to acquire, speak and comprehend language as a natural, biologically-inherited one. One characteristic of work in semiotic accounts of language and communication is to stress its *un*natural aspects, focussing on the socially-regulated, arbitrary, conventional nature of meanings.

Seen in the light of Grice's myth, however, the unnatural '[*arbitrary/conventional*] symbol-using mind' (Vol. 2, p. 299)[42] of Peirce and the

[42] Quote from Feibelman's '*An Introduction to Peirce's Philosophy*' (pp. 91-92); page reference to Peirce's '*Collected Writings' Volume 1-6.*

semioticians can be re-interpreted. Once we have reached a stage where words no longer have to be 'natural signs of that they are used to communicate', we have reached a stage not only where individuals have the ability to reflect on the intentions behind instances of their use, but also to reflect *on the content of the signals themselves*. Indeed, it could be argued that the two abilities are fundamentally the same. Construed in this way, *non*-natural doesn't mean "*un*natural", and that, I think, means that we're probably on the right track.

Bibliography

Aitchison, J. (1991) *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.

Aitchison, J. (1996) *The Seeds of Speech.* Cambridge: Cambridge University Press.

Akhtar, N., M. Carpenter and M. Tomasello (1996) The role of discourse novelty in children's early word learning. *Child Development, 67*; 635-645.

Allen, C. and M. Bekoff (1997) *Species of Mind.* Cambridge, MA: MIT Press.

Ameka, F. (1992) Interjections: the universal yet neglected part of speech. In *Journal of Pragmatics 18*; 101-118.

Arundale, R. (1991) Studies in the way of words: Grice's new directions in conceptualizing meaning in conversational interaction. Paper presented to the *International Communication Association*: Chicago, IL. May 23rd – 27th, 1991.

Avramides, A. 1989: *Meaning and Mind: An Examination of a Gricean Account of Language.* Cambridge, MA: MIT Press.

Austin, J. (1962) *How to Do Things with Words.* Oxford: Clarendon Press.

Bach, K. (1994) Conversational Impliciture. *Mind and Language 9*; 124-162.

Bach, K. and R. Harnish (1979) *Linguistic Communication and Speech Acts.* Cambridge, MA: MIT Press.

Banfield, A. (1982) *Unspeakable Sentences: Narration and Representation in the Language of Fiction.* Boston, MA: Routledge and Kegan Paul.

Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind.* Cambridge, MA: MIT Press.

Baron-Cohen, S., A. Spitz and P. Cross (1993) Do children with autism recognise surprise? A research note. *Cognition and Emotion. 7(6)*; 507-516.

Barwise, J. (1989) *The Situation in Logic.* Centre for the Study of Language and Information—distributed by University of Chicago Press.

Bateson, M-C. (1996) On the naturalness of things. In Brockman, J. and K. Matson (eds.) *How Things Are: A Science Tool-kit for the Mind.* Phoenix: London.

Benfey, T. (1869) *Geschicte der Sprachwissenschaft.* Munich.

Bickerton, D. (2000) Calls aren't words, syllables aren't syntax. *Psycoloquy 11*, No. 14.

Blake, J., S. McConnell, G. Horton and N. Benson (1992) The gestural repertoire and its evolution over the second year. *Early Development and Parenting, 1*; 127-136.

Blakemore, D. (1987) *Semantic Constraints on Relevance.* Oxford: Blackwell.

Blakemore, D. (1992) *Understanding Utterances: An Introduction to Pragmatics.* Oxford: Blackwell.

Blakemore, D. (2002) *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers.* Cambridge: Cambridge University Press.

Blass, R. (1990) *Relevance Relations in Discourse: A Study with Special Reference to Sissala.* Cambridge: Cambridge University Press.

Bloom, P. and T. German (2000) Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*; B25-B31.

Bolinger, D. (1983) The inherent iconism of intonation. In Haiman, J. (ed.) (1983) *Iconicity in Syntax (Typological studies in language, Volume 6).* Amsterdam/Philadelphia: John Benjamins; 97-109

Bradbury, J. and S. Vehrencamp (1998) *Principles of Animal Communication.* Sunderland, MA: Sinauer Associates.

Breheny, R. (forthcoming) Communication and theory of mind. *Ms.* University of Cambridge.

Bretherton, I. (1991) Intentional communication and the development of an understanding of mind. In Frye, D. and C. Moore (eds.) *Children's Theories of Mind: Mental States and Social Understanding.* Hillsdale, NJ: Lawrence Erlbaum; 49-75.

Byrne, R. and A. Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans.* Oxford: Clarendon Press.

Carstairs-McCarthy, A. (1999) *Origins of Complex Language.* Oxford: Oxford University Press.

Carston, R. (1996) Enrichment and loosening: complimentary processes in deriving the proposition expressed. *UCL Working Papers in Linguistics 8*; 61-88.

Carston, R. (2000) Explicature and semantics. *UCL Working Papers in Linguistics 12*; 1-44.

Carston, R. (2002*a*) *Thoughts and Utterances: The Pragmatics of Explicit Communication.* Oxford: Blackwell.

Carston, R. (2002*b*) Linguistic meaning, communicated meaning and pragmatics. *Mind and Language, Vol. 117, Nos. 1 and 2.* Oxford: Blackwell; 127-148.

Cheney D. and R. Seyfarth (1990) *How Monkeys See the World: Inside the Mind of Another Species.* Chicago: Chicago University Press.

Chomsky, N. (1986) *Knowledge of Language.* London: Praeger.

Chomsky, N. (1988) *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (1999) An Online Interview with Noam Chomsky: on pragmatics and related issues. *Brain and Language, Vol. 68, No. 3*. Academic Press; 393-401.

Chomsky, N. (2000) *New Horizons in the Study of Language and Mind.* Cambridge: Cambridge University Press.

Clark, B. and R. Gerrig (1990). Quotations as demonstrations. *Language 66*; 764-805.

Clark, H. (1994) *Using Language*. Cambridge: Cambridge University Press.

Cole, P. and J. Morgan (eds.) (1975) *Syntax and Semantics 3: Speech Acts*. New York, Academic Press.

Cosmides, L. (1989) The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition 31*; 197-276.

Crystal, D. (1995) *The Cambridge Encyclopaedia of Linguistics*. Cambridge: Cambridge University Press.

Cuxac, C. (1999) The expression of spatial relations and the spatialization of semantics in French Sign Language. In Fuchs, C. and S. Roberts (eds.) *Language Diversity and Cognitive Representations*. Amsterdam: Benjamins; 123-142.

Davis, S. (1991) *Pragmatics: A Reader*. Oxford: Oxford University Press.

Dale, R. (1996) *Theory of Meaning*. PhD Thesis. City University of new York. Available online at http://russelldale.com/dissertation/parts.html.

Darwin, C. (1872) *The Expression of the Emotions in Man and Animals*. London: John Murray. Republished (1998) (Edited with foreword, commentary and afterword by P. Ekman). London: Harper Collins.

Davidson, D. and G. Harman (eds.) (1975) *The Logic of Grammar*. Dordrecht: Reidel.

Davies, M. (1996) Philosophy of language. In: Bunnin, N. and E.P Tsui-James (eds.) *The Blackwell Companion to Philosophy*. Oxford: Blackwell; 90-139.

Deacon, T. (1997) *The Symbolic Species*. London: Penguin.

Demko, T. (1995) Which came first: bees or flowers? *Science Times*; *The New York Times*. Tuesday 23rd May 1995; B7-B8.

Dennett, D. (1978) Beliefs about beliefs. *Behavior and Brain Sciences, 4*; 568-570.

Dennett, D. (1987) *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D. (1988) The Intentional Stance in Theory and in Practice. In Byrne, R. and A. Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and The Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon Press; 180-202.

Dennett, D. (1995) *Darwin's Dangerous Idea - Evolution and the Meanings of Life*. London: Penguin.

Dennett, D. (1996) Granny versus Mother Nature—No Contest. *Mind and Language 11*; 263-269.

Donald, M. (1998) Mimesis and the executive suite. In Hurford, J., Studdert-Kennedy, M and C. Knight (eds.) *Approaches to the Evolution of Language*. Cambridge: Cambridge University Press; 44-68.

Dreller, C. and W. Kirchner (1993) How honeybees perceive the information of the dance language. *Naturwissenshaften, 80*; 319-321.

Dreller, C. and W. Kirchner (1994) Hearing in the Asian honeybees *Apis dorsata* and *Apis florae. Insectes Socieaux, 42*; 115-125.

Duchenne, G. (1862) *Mecanisme de la Physionomie Humaine* Republished as (1990) *The Mechanism of Human Facial Expression* (Edited and translated by R. A. Cuthbertson) Cambridge: Cambridge University Press.

Dummett, M. (1981) *The Interpretation of Frege's Philosophy.* London: Duckworth.

Dunbar, R. (1998) Theory of mind and the evolution of language. In Hurford, J., Studdert-Kennedy, M and C. Knight (eds.) *Approaches to the Evolution of Language.* Cambridge: Cambridge University Press; 92-111.

Eibl-Eibesfeldt, I. (1972) Similarities and differences between cultures in expressive movements. In Hinde, R. (ed.) *Non-Verbal Communication.* Cambridge: Cambridge University Press; 297-312.

Efron, D. (1941) *Gesture and Environment.* Morningside Heights, New York: King's Crown Press.

Ekman, P. and W. Friesen (1969) The repertoire of non-verbal behaviour categories: Origins, usage and coding. *Semiotica 1*; 49-98.

Ekman and Rosenberg (1997) *What the Face Reveals.* New York: Oxford University Press.

Ekman, P. (1989) The argument and evidence about universals in facial expressions of emotion. In Wagner, H. and A. Manstead (eds.) *Handbook of Social Psychophysiology.* New York: Wiley; 143-164.

Ekman, P. (1992) An argument for basic emotion. *Cognition and Emotion 6 (3/4)*; 169-200.

Ekman, P. (1994) Strong evidence for universals in facial expressions: A Reply to Russell's Mistaken Critique. *Psychological Bulletin 115 (2)*; 268-287.

Ekman, P. (1999) Emotional and conversational nonverbal signals. In Messing, L. and R. Campbell (eds.) *Gesture, Speech and Sign*. Oxford: Oxford University Press; 45-57.

Feibelman, J. (1960) *An Introduction to Peirce's Philosophy*. London: Allen and Unwin.

Fodor, J., Fodor, J. and M. Garrett (1975) The psychological unreality of semantic representations. In *Linguistic Enquiry, Fall 1975*; 515-531.

Fodor, J. (1975) *The Language of Thought*. New York: Crowell.

Fodor, J. (1981) *Representations*. Hassocks: Harvester Press.

Fodor, J. (1983) *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. (1998) *Concepts: Where Cognitive Science Went Wrong*. Oxford: Clarendon Press.

Fridlund, A. (1994) *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press.

von Frisch, K. (1967) *The Dance Language and Orientation of Bees*. Cambridge, MA: Belknap Press of Harvard University Press.

Gardiner, A. (1951). *The Theory of Speech and Language*. 2nd edition. Oxford: Clarendon Press. (First published: Oxford, 1932.)

Gazzaniga, M. and C. Smiley (1991) Hemispheric mechanisms controlling voluntary and spontaneous facial expressions. *Journal of Cognitive Neuroscience 2*; 239-245.

Gigerenzer, G. and P. Todd (1999) Fast and frugal heuristics: the adaptive toolbox. In Gigerenzer, G., P. Todd and the ABC Research Group. *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Gigerenzer, G., P. Todd and the ABC Research Group. *Simple Heuristics That Make Us Smart*. Oxford: Oxford University Press.

Goffman, E. (1964) The neglected situation. *American Anthropologist*, Vol. 66, No. 6, Part 2; 133-136.

Goffman, E. (1981) *Forms of talk*. Oxford: Blackwell.

Goodglass, H. (1993) *Understanding Aphasia*. New York: Academic Press.

Gould, J. (1990) Honeybee cognition. *Cognition, 37*; 83-103.

Gould, S. J. (1989) Tires to Sandals. In *Natural History*, April 1989, 8-15.

Gould, S. J. and R. Lewontin (1979) The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program. In Selzer, J. (1993) *Understanding Scientific Prose*. Wisconsin: University of Wisconsin Press.

Grandy, R. and R. Warner (1986) Paul Grice: a view of his work. In Grandy, R. and R. Warner (eds.) *Philosophical Grounds of Rationality: Intentions, Categories, Ends*. Oxford: Clarendon Press; 1-44.

Grice, H. P. (1957) Meaning. In *Philosophical Review 66*; 377-88.

Grice, H. P. (1967) *William James Lectures I-VII*. Unpublished typescripts.

Grice, H. P. (1968) Utterer's meaning, sentence meaning and word-meaning. *Foundations of Language 4*; 225-242.

Grice, H. P. (1969) Utterer's meaning and intentions. *Philosophical Review 78*; 147-177.

Grice, H. P. (1975) Logic and conversation. In Cole, P. and J. Morgan (eds.) *Syntax and Semantics 3: Speech Acts*. New York: Academic Press; 41-58.

Grice, H. P. (1982) Meaning revisited. In Smith, N. (ed.) *Mutual Knowledge*. London: Academic Press.

Grice, H. P. (1989) *Studies In The Way of Words*. Cambridge, MA: Harvard University Press.

Grice, H. P. (2001) *Aspects of Reason*. (edited by Warner, R.) Oxford: Clarendon Press

Gussenhoven, C. (2002) Intonation and interpretation: Phonetics and Phonology. Speech Prosody 2002: *Proceedings of the First International Conference on Speech Prosody*. Aix-en-Provence, ProSig and Universite de Provence Laboratoire Parole et Language. 47-57.

Hamilton, E. and H. Cairns (1989) *The Collected Dialogues of Plato*. Princeton, NJ: Princeton University Press.

Happé, F. (1993) Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition 48*; 101-19.

Happé, F. (1994) *Autism: An Introduction to Psychological Theory*. Cambridge MA: Harvard University Press.

Happé, F. (1992) The autobiographical writings of three Asperger syndrome adults: Problems of interpretation and implications for theory. In Frith, U. (ed.) *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press; 207-242.

Happé, F. and E. Loth. (2002) Theory of mind and tracking speakers' intentions. in *Mind and Language, Vol. 117, Nos. 1 and 2*. Oxford: Blackwell; 24-36.

Hare, R. M. (1949) Imperative sentences. *Mind, LVIII.* Republished in Hare, R. M. (1971) *Practical Inferences.* London: Macmillan; 1-21.

Hare, R. M. (1970) Meaning and speech acts. *Philosophical Review, LXXIX.* Republished in Hare, R. M. (1971) *Practical Inferences.* London: Macmillan; 74-93.

Harris, J. and E. Gussman (2002) Word-final onsets. *UCL Working Papers in Linguistics 14*; 1-42.

Hart, H. L. A. (1952) Signs and words. In *Philosophical Quarterly Vol. 2.* The University of St. Andrews for The Scots Philosophy Club.

Hauser, M. (1996) *The Evolution of Communication.* Cambridge, MA: MIT Press.

Hauser, M., N. Chomsky, and T. Fitch (2002) The faculty of language: what is it? Who has it, and how did it evolve? *Science 298*; 1569-1579.

Hobson, P.; Ouston, J. and A. Lee (1988) What's in a face? The case of autism. *British Journal of Psychology. Vol. 79, (4)*; 441-453

Hockett, C. (1959) Animal 'languages' and human language. In J. N. Spuhler (ed.) *The Evolution of Man's Capacity for Culture.* Detroit: Wayne State University Press; 32-39.

Holloway, J. (1951) *Language and Intelligence.* London: Macmillan.

van Hooff, J. (1972) A comparative approach to the phylogeny of laughter and smiling. In Hinde, R. (ed.) *Non-verbal Communication.* Cambridge: Cambridge University Press; 209-238.

Humphrey, N. (1988) The Social Function of Intellect. In Byrne, R. and A. Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and The*

Evolution of Intellect in Monkeys, Apes and Humans. Oxford: Clarendon Press; 271-286.

Ifantidou-Trouki, E. (1993) Sentential adverbs and relevance. In *Lingua* 90 (1); 65-90.

Ifantidou, E. (2001) *Evidentials and Relevance.* Amsterdam: John Benjamins.

Itani, R. (1995) *Semantics and Pragmatics of Hedges in English and Japanese.* University of London. PhD Thesis.

Iten, C. (2000) 'Non-truth-Conditional' Meaning, Relevance and Concessives. University of London PhD Thesis.

Jakobson, R. and M. Halle (1956) *Fundamentals of Language.* The Hague: Mouton.

Jespersen, O. (1922) *Language: Its Nature, Development and Origin.* London: Allen and Unwin.

Kaplan, D. (1977) Demonstratives. In J. Almog, J. Perry and H. Wettstein (eds.) (1989) *Themes from Kaplan.* Oxford: Oxford University Press; 481-563.

Kaplan, D. (1997) What is meaning? Explorations in the theory of meaning as use. *Ms.*

Kazantzakis, N. (1965) *Travels in Greece (Journey to the Morea).* Translated by F. A. Reed. Oxford: Bruno Cassirer.

Kendon, A. (1988) How gestures can become like words? In Poyotas, F. (ed.) *Cross-cultural Perspectives in Nonverbal Communication.* Toronto: Hogrefe; 131-141.

Kraut, R. and R. Johnston (1979) Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology, 42*; 1529-1553.

Lass, R. (1997) *Historical Linguistics and Language Change.* Cambridge: Cambridge University Press.

Leslie, A. (1987) Pretense and representation: The origins of 'theory of mind'. *Psychological Review 94*; 412-426.

Leslie, A. (1994) ToMM, ToBY and Agency: core architecture and domain-specificity. In Hirschfeld L. and S. Gelman (eds.) *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge: Cambridge University Press; 119-148.

Leslie, A. and F. Happé (1989) Autism and ostensive communication: The relevance of metarepresentation. *Development and Psychopathology 1*; 205-212.

Levinson, S. (1989) *Pragmatics.* Cambridge: Cambridge University Press.

Lewis, D. (1969) *Convention.* Cambridge, MA: Harvard University Press.

Lightfoot, D. (1991) *How to Set Parameters: Arguments from Language Change.* Cambridge, MA: MIT Press.

Lycan, W. (1991) Review of Avramides, A. (1989) *Meaning and Mind: An Examination of a Gricean Account of Language.* In *Mind and Language, 6*; 83-86.

Lyons, J. (1972) Human Language. In Hinde, R. (ed.) *Non-Verbal Communication.* Cambridge: Cambridge University Press; 49-86.

McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

Miller, G. (1978) Semantic relations among words. In Halle, M., J. Bresnan and G. Miller (eds.) *Linguistic Theory and Psychological Reality.* Cambridge, MA: MIT Press; 60-118.

Millikan, R. (1984) *Language, Thought and Other Biological Categories.* Cambridge, MA: MIT Press.

Moles, A. (1963) Animal language and information theory [Originally published as 'Langage animal et théorie de l'information'] In Busnel, R. (ed.) *Acoustic Behaviour of Animals.* Amsterdam: Elsevier; 112-131.

Morris, C. H. (1938) Foundation of the theory of signs. Reprinted in Morris, C. (1971) *Writings on the General Theory of Signs.* The Hague: Mouton; 13-71.

Muller, M. (1862) *Lectures on the Science of Language.* New York: Charles Scribner.

Muris, P., Meesters, C., Merckelbach, H. and M. Lomme (1995) Knowledge of basic emotions in adolescent and adult individuals with autism. *Psychological Reports. Vol. 76(1), Feb 1995*; 52-54.

Neale, S. (1992) Paul Grice and the philosophy of language. *Linguistics and Philosophy 15.5*; 509-559.

Nuti, M. (1999) On folk-psychology. *University College London Working Papers in Linguistics, Volume 11*; 81-101.

Ogden, C. and I. Richards (1923) *The Meaning of Meaning: A Study of the Influence of Language upon Thought.* London: Routledge and Kegan Paul.

Okanoya, K. (forthcoming) Sexual selection can shape finite-state syntax in birdsong: a sexual display origin of syntax. *Ms.* Chiba University.

Origgi, G. and D. Sperber (2000) Evolution, communication and the proper function of language. In Carruthers, P. and A. Chamberlain (eds.), *Evolution and the Human Mind: Modularity, Language and Meta-cognition.* New York: Cambridge University Press; 140-169.

Ozonoff, S., Pennington, B. and S. Rogers (1990) Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry and Allied Disciplines. Vol. 31 (3), Mar 1990*; 343-361.

Peirce, C. (1897, 1903) Logic as semiotic: the theory of signs. In J. Buchler (ed.) (1955) *The Philosophical Writings of Peirce*. New York: Dover Press.

Pinker, S. and P. Bloom (1990) Natural Language and Natural Selection. In *Behavioural and Brain Sciences 13*; 707-784.

Quirk, R., Greenbaum, S., Leech, J. and J. Svartik (eds.) (1985) *A Comprehensive Grammar of the English Language*. London: Longman.

Recanati, F. (1987) *Meaning and Force: The Pragmatics of Performative Utterances*. Cambridge: Cambridge University Press.

Recanati, F. (1993) *Direct Reference*. Oxford: Blackwell.

Recanati, F. (1998) Pragmatics. In *Routledge Encyclopaedia of Philosophy*, Vol. 7. London: Routledge; 620-633.

Recanati, F. (2002) Does linguistic communication rest on inference? *Mind and Language, Vol. 117, Nos. 1 and 2*. Oxford: Blackwell; 105-126.

Recanati, F. (forthcoming) Unarticulated constituents. *Linguistics and Philosophy*.

Rey, G. (1980) Functionalism and the emotions. In Rorty, A. (ed.) *Explaining Emotions*. Los Angeles: University of California Press; 163-198.

Rinn, W. (1984) The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, Volume 95*; 52-77.

van Rooten, H. K. (1968) *Mots d'Heures: Gousses, Rames*. Brighton: Angus and Robertson.

Russell, J. (1994) Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. *Psychological Bulletin 115*; 102-141.

Sacks, O. (1994) A neurologist's notebook: An Anthropologist on Mars. *New Yorker*, December 27th 1993 – January 3rd 1994.

Sadock, J. and A. Zwicky (1985) Speech act distinctions in syntax. In Shopen, T. (ed.) *Language Typology and Syntactic Description. Volume 1.* Cambridge: Cambridge University Press; 155-196.

Sapir, E. (1970) *Language: An Introduction to the Study of Speech.* London: Harvest.

Saul, J. (2002) What is said and psychological reality: Grice's project and relevance theorists' criticisms. *Linguistics and Philosophy, 25*; 347-372.

Saussure, F. de (1974) *Course in General Linguistics.* (First published 1916). Peter Owen: London.

Schiffer, S. (1972) *Meaning.* Oxford: Clarendon Press.

Schiffrin, B. (1987) *Discourse Markers.* Cambridge: Cambridge University Press.

Scholl, B. and A. Leslie (1999) Modularity, development and 'theory of mind'. *Mind and Language 14*; 131-153.

Searle, J. (1965) What is a speech act? In M. Black (ed.) *Philosophy in America.* Allen and Unwin; 221-239.

Searle, J. (1969) *Speech acts.* Cambridge: Cambridge University Press.

Searle, J. (1979) *Expression and Meaning.* Cambridge: Cambridge University Press.

Sebeok, T. (1972) *Perspectives in Zoosemiotics.* The Hague: Mouton.

Seyfarth, R. and D. Cheney (1992) Meaning and mind in monkeys. In *Scientific American. Vol. 267/6*; 78-86.

Seyfarth, R., Cheney, D. and P. Marler (1980) Vervet monkey alarm calls: semantic communication in a free-ranging primate. In *Animal Behaviour 28*; 1070-1094.

Shannon, C. and W. Weaver (1949) *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Smith, C. and H. Scott (1997) A componential approach to the meaning of facial expressions. In Russell, J. and J. Fernandez-Dols (eds.) *The Psychology of Facial Expression.* New York: Cambridge University Press;

Smith, N. and I. Tsimpli (1995) *The Mind of a Savant*. Oxford: Blackwell.

Sperber, D. (1990) The Evolution of the language faculty: A paradox and its solution. In *Behavioural and Brain Sciences 13*; 756-758.

Sperber, D. (1994) The modularity of thought and the epidemiology of representations. In Hirschfeld, L. and S. Gelman (eds.) *Mapping the Mind: Domain Specificity in Cognition and Culture.* Cambridge: Cambridge University Press; 39-67.

Sperber, D. (1995) How do we communicate? In Brockman, J. and K. Matson (eds.) *How Things Are: A Science Tool-kit for the Mind.* Phoenix: London; 191-199.

Sperber, D. (2000) Metarepresentations in an evolutionary perspective. In D. Sperber (ed.) *Metarepresentations : A Multidisciplinary Perspective.* Oxford: Oxford University Press; 117-137.

Sperber, D. (2001) An evolutionary perspective on testimony and argumentation. To appear in a special issue of *Philosophical Topics* in honour of Alvin Goldman.

Sperber, D. and D. Wilson (1986/1995) *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D and D. Wilson (1998) The mapping between the mental and public lexicon. In Carruther, P. and J. Bouchers (eds.) *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press; 184-200.

Sperber, D and D. Wilson (2002) Pragmatics, modularity and mindreading. In *Mind and Language, Vol. 17*. Oxford: Blackwell; 3-23.

Stanley, J. (2000) Context and logical form. *Linguistics and Philosophy 23*; 391-434.

Stevenson, C. (1944). *Ethics and Language*. New Haven and London: Yale University Press.

Strawson, P. (1964) Intention and convention in speech acts. *Philosophical Review 73*; 439-460.

Suppes, P. (1996) The primacy of utterer's meaning. In Grandy, R. and R. Warner (eds.) *Philosophical Grounds of Rationality: Intentions, Categories, Ends*. Oxford: Clarendon Press; 109-129.

Tomasello, M. and M. Barton (1996) Learning words in non-ostensive contexts. *Cognitive Development, 10*; 201-224.

Trask, R. L. (1993) A Dictionary of Grammatical Terms in Linguistics. London: Routledge.

Vygotsky, L. (1962) *Thought and Language*. (Translated by E. Hanfmann and G. Vakar) Cambridge, MA: MIT Press.

Welby, Lady Victoria (1893) Meaning and metaphor. *Monist, 3:4 (1893)*; 510-525. Reprinted in Welby (1985).

Welby, Lady Victoria (1896) Sense, meaning and interpretation. *Mind, 5:17* (1896); 24-37 and *Mind, 5:18* (1896). Reprinted in Welby (1985).

Welby, Lady Victoria (1911) Reprinted as Welby (1985) H. Walter Schmitz (ed.) *Significs and Language: The Articulate Form of Our Expressive and Interpretive Resources*. Amsterdam and Philadelphia: John Benjamins.

Wharton, T. (2000*a*) Interjections, evolution and the 'showing'/'saying' continuum. Paper presented at *The Third International Conference on the Evolution of Language (Evolang)*. Paris, 3rd – 6th April, 2000.

Wharton, T. (2000*b*) Interjections, language and the 'showing'/'saying' continuum. *UCL Working Papers in Linguistics 12*; 173-213.

Wharton T. (2001) Natural pragmatics and natural codes. *UCL Working Papers in Linguistics 13*; 109-158.

Wharton, T. (2002) Paul Grice, saying and meaning. *UCL Working Papers in Linguistics 14*; 207-248

Wharton, T. (2003*a*) Interjections, language and the 'showing'/'saying' continuum. To appear in *Pragmatics and Cognition. Vol. 11-1*. May 2003.

Wharton, T. (2003*b*) Natural pragmatics and natural codes. To appear in *Mind and Language*. September 2003.

Whiten, A. and R. Byrne (1988) Taking (Machiavellian) Intelligence Apart: editorial. In Byrne, R. and A. Whiten (eds.) (1988) *Machiavellian Intelligence: Social Expertise and The Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon Press; 50-66.

Whiten, A. and R. Byrne (1997) *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge: Cambridge University Press

Wierzbicka, A. (1992) The semantics of interjection. In *Journal of Pragmatics 18*; 159-192.

Wierzbicka, A. (1994) Cognitive domains and the structure of the lexicon: The case of emotions. In Hirschfeld L. and S. Gelman (eds.) *Mapping the Mind: Domain Specificity in Cognition and Culture*. Cambridge: Cambridge University Press; 431-453.

Wierzbicka, A. (1996) *Semantics: Primes and Universals*. Oxford: Oxford University Press.

Wierzbicka, A. (2000) The semantics of human facial expression. In *Pragmatics and Cognition 8 (1)*; 147-183.

Wilkins, D. (1992) Interjections as deictics. In *Journal of Pragmatics 18*; 119-158.

Williams, D. (1992) *Nobody Nowhere: The Remarkable Autobiography of an Autistic Girl.* London: Jessica Kingsley.

Williams, D. (1994) *Somebody Somewhere.* London: Jessica Kingsley.

Williams, D. (1999) *Like Colour to the Blind.* London: Jessica Kingsley.

Wilson, D. (1998) Linguistic Structure and Inferential Communication. In Caron, B. (ed.) *Proceedings of the 16th International Congress of Linguists*. Oxford: Pergamon.

Wilson, D. (2000) Metarepresentation in linguistic communication. In D. Sperber (ed.) *Metarepresentations : A Multidisciplinary Perspective.* Oxford: Oxford University Press; 411-448.

Wilson, D. and D. Sperber (1981) On Grice's theory of conversation. In Werth, P. (ed.) (1981) *Conversation and discourse*. Croom Helm, London.

Wilson, D. and D. Sperber (1988) Mood and the analysis of non-declarative sentences. In Dancy, J., Moravcsik, J. and C. Taylor (eds.) *Human Agency: Language, Duty and Value*. Stanford University Press. Stanford: CA; 77-101.

Wilson, D. and D. Sperber (1993) Linguistic form and relevance. In *Lingua 90 (1)*; 1-25.

Wilson, D. and D. Sperber (2002) Truthfulness and Relevance. In *Mind, Vol. 111-442* July 2002; 583-632.

'There is a correspondence between the microcosm and the macrocosm! The stars are alive, child. Did you know that? Everything out there is alive, and there are grand purposes abroad! The universe is full of *intentions*, you know.'

(Philip Pullman—*Northern Lights*)