Natural Pragmatics and Natural Codes[*]

TIM WHARTON

Abstract

Grice (1957) drew a famous distinction between natural$_{(N)}$ and non-natural$_{(NN)}$ meaning, where what is meant$_{(NN)}$ is broadly equivalent to what is intentionally communicated. This paper argues that Grice's dichotomy overlooks the fact that spontaneously occurring natural signs may be intentionally *shown*, and hence used in intentional communication. It also argues that some naturally occurring behaviours have a signalling function, and that the existence of such *natural codes* provides further evidence that Grice's original distinction was not exhaustive. The question of what kind of information, in cognitive terms, these signals encode is also examined.

1 Introduction

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 these behaviours. The linguist seeks to sift out extraneous, non-linguistic phenomena, and focus on the rule-based grammar—the *code*. 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, 1995), and are now in a position to ask questions it was once not even possible to formulate (Chomsky 2000).

[*]I would like to thank Deirdre Wilson and Robyn Carston, both for their continued patience and support, and for providing numerous insightful comments on previous drafts of this paper. I would also like to thank The UCL Graduate School, without whose funding I would be unable to continue my studies.

There are two reasons, however, why the pragmatist should cast a broader net. Firstly, thanks largely to the influential work of Paul Grice (1957, 1969, 1975[1]), it is now increasingly recognised that verbal comprehension is more than a simple coding-decoding process. Any attempt to characterise linguistic *communication* should reflect the fact that it is an intelligent, intentional[2] activity. Secondly, the aim of a pragmatic theory is to explain how *utterances* are understood; the task, therefore, of describing and explaining precisely what certain natural behaviours indicate, and how they are interpreted, would appear to fall squarely within the domain of pragmatics.

In this paper I focus on the following questions:

(a) What is the relationship between the natural behaviours described above and intentional communication?
(b) What do these natural behaviours indicate?
(c) How are they interpreted?

In Section 2 I address question (a). I discuss Grice's seminal paper 'Meaning' (1957), and argue that the distinction between intentional communication and other forms of information transmission cross-cuts Grice's famous distinction between non-natural and natural meaning. While it is clear that certain involuntary/natural indicators *betray* our thoughts and feelings to others in a way we would not want to describe as intentionally communicating them, I will argue that some are used intentionally: they

---

[1] Page references to these works are from Grice's 1989 anthology *Studies in the Way of Words*.

[2] I use the word 'intentional' in what Hauser (1998: 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 in what follows 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 & Bekoff 1997), concerned as they are with mapping the abilities of non-human animals to attribute mental states to others (although the notion Allen & 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 paper). One more point on terminology: the abilities that govern the attribution of mental states are now increasingly referred to as 'metarepresentational' abilities (see Sperber 2000). I stick with '*intentional* communication' here, at least to begin with, because my route into the discussion is via Grice's original paper.

may not be deliberately *produced*, but that does not mean they cannot be deliberately (or intentionally) *shown*.[3]

In Section 3 I present the pragmatic framework that I adopt in this paper—Relevance Theory (Sperber & Wilson 1986, 1995). This is a cognitive-inferential model, which builds on Gricean foundations, and provides the framework within which I approach question (b). In Section 4 I address question (c) and argue that natural behaviours do not all work in the same way. Many things in the world carry information, or 'indicate': tree-rings, footprints in snow, the scent of ripe fruit. 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 natural behaviours too—behaviours that Grice regarded as natural signs. 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 are best analysed as ***natural codes***. The existence of these inherently communicative natural signs provides further evidence that Grice's original distinction is not exhaustive.

The existence of natural codes leads me to pose a fourth question ((d) below):

(d) In cognitive terms, what kind of information do human natural codes encode?

To answer this question I first attempt to clarify the notion *code* itself, and then look at the strengths and weaknesses of two previous attempts to describe and explain natural codes—those of Thomas A. Sebeok (1972) and Anna Wierzbicka (2000). Finally, I present my own account of the coded element in natural codes, which preserves the insights of these two previous approaches.

2 Meaning and *intentional* communication

2.1 Meaning$_{N}$ and meaning$_{NN}$

In his famous paper 'Meaning' (1957), Paul Grice proposed a distinction between *natural*$_{(N)}$ and *non-natural*$_{(NN)}$ meaning. This was intended to reflect what he later

[3] Although this paper focuses largely on 'intentionally' used natural behaviours, this is not to suggest that those that are perhaps beyond its scope—pupil dilation, pheromones—should not be integrated with such a model.

described 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: 291).

Cases of each kind are given in (1) and (2) below:

(1) Those spots mean[4] measles.
(2) That remark means he has measles.

Grice proposed a variety of ways in which the two types of meaning might be distinguished. 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, *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. This observation was reflected in one of a series of tests, in which Grice contrasted ways that 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.

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

Another recognition 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".

---

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

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.

Of the two types of meaning, Grice was primarily concerned with meaning$_{NN}$; in particular, how the kind of meaning exemplified in (4b) and (5b) might be characterised in terms of intentions and the recognition of intentions. Consider (6a-e) below—adapted from Grice (1989: 218). In all these sentences something has happened that has produced a response in an audience. In both senses of the word, then, Mary's pale complexion and her utterance might be said to *mean* she is unwell:

(6a) Mary is asleep. Her mother notices that she is pale and concludes she is unwell.
(6b) Feeling unwell, Mary intends her mother to see how pale she is. However, she doesn't want this intention to be noticed (it might decrease her chances of getting a day off school), so she pretends to be asleep.
(6c) As (6b), except that here Mary's mother realises Mary is pretending to be asleep, and hence infers her intention not to have her intention noticed.
(6d) Feeling unwell, Mary deliberately and openly lets her mother see how pale she is, so she will notice and help.[5]

[5] 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 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 open about their intentions (the 'handkerchief at the scene of the murder' example (1989: 217)), it seems clear that what Grice had in mind in his original example was a case in which the child (Mary in my adaptation) 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: 218—and see p. 6 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 deep philosophical discussions (and very quickly out of my depth) about whether 'hoping'

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

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 have a true (Gricean) *communicative* intention: the intention that the recognition of her intention to produce the desired response will play a crucial role in producing the response itself.[6]

Although it will be necessary to look at it in a little more detail later, it is not the aim of this paper to discuss Grice's characterisation of meaning$_{NN}$. Instead, I will focus largely on the notion of meaning$_{N}$. The essence of my argument will be that the distinction between non-intentional and intentional communication crosscuts Grice's natural/non-natural dichotomy.

This is not intended as a criticism of Grice, who did not consider meaning$_{N}$ in communicative terms. To him, cases of meaning$_{NN}$ just are those "cases which are related to communication" (1989: 291), and cases of meaning$_{N}$ are instances of what Davies (1996: 116) calls "*indicator* meaning", or mere "causal co-variation between two kinds of states of affairs, which indicate whether or not anyone takes them in that way, and whether or not anyone intends them to be taken in that way". Seen in this way, Mary's pale complexion means$_{N}$ she is unwell independent of any intentions Mary may have in deliberately and openly showing her mother her pale complexion.

Throughout his work, Grice saw the link between meaning and intention as a point from which he might begin an exploration of *word* meaning and *sentence* meaning: Davies again—"the programme begun by Grice aims at an analysis of the concept of *literal meaning in a public language*…" (1996: 115 (my emphasis, TW)). For this

---

necessarily involves 'intending': it certainly doesn't always – I think you can hope for something without intending it; however, you can equally 'hope' *and* 'intend' something. That is, you can hope that *P*, and intend that *P* too.

[6] Grice (1957) does not use the phrase 'communicative intention'. I use it (following Sperber & Wilson (1986/1995)) to distinguish the intention to communicate from the intention to inform (which Mary clearly has in (6b-e)). Grice (1968/1989: 123) abbreviates his original 3-clause definition with the phrase *'M-intend'*.

reason, he was keen to draw a line between cases such as (6d) and (6e) above, between "'deliberately and openly letting someone know' and 'telling'"[7] (1989: 218).

Was Grice right to draw this line? Given his aims, we should not be critical. However, the goals of pragmatic theory extend beyond an exploration of semantic notions such as those Grice (1957) had in mind, to providing an account of human intentional communication generally. This shift in emphasis, I argue, requires that we draw the line between intentional communication and other forms of information transmission in a different place.

2.2 Intentional communication

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. Rather, it is 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 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.* During a presentation at a conference, a speaker might intend the audience to form the impression that he is not at all nervous. The successful fulfilment of this intention depends entirely on its remaining concealed, for the audience's discovery that the speaker has this intention will usually be enough for them to conclude that actually the speaker is nervous after all.

But what of the cases in (6d) and (6e)? Does the line that Grice drew reflect a distinction between intentional communication and other forms of information transmission? If we equate 'intentionally communicate' with meaning$_{NN}$, it is clear that it does. However, 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 there are good reasons to suggest that (6d) might indeed qualify as an instance of intentional communication, though this concept will have to be distinguished from meaning$_{NN}$.

---

[7] Though Grice did not restrict meaning$_{NN}$ to *linguistic* acts.

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 Mary's pale complexion showing her mother she is unwell (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.

In an act of overt communication, then, a communicator must provide evidence of their intention to inform. In any act carried out with the intention of providing evidence of an *informative* intention there are two layers to be retrieved. The first, basic layer is the information being pointed out—in (6d) the fact that Mary is unwell and wants help—and the second is the information that the first layer is being pointed out intentionally. Grice's communicative intention was formulated in such a way as to suggest that this basic layer should not be entirely derivable without reference to the second. In the light of the above arguments, we might say that for a communicative act to be *intentional*, 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.

The Gricean communicative intention has been the subject of a great deal of discussion (Strawson 1964, Searle 1969, Schiffer 1972), and numerous revisions. The point being made here is that by excluding cases that should be included as instances of intentional communication, it is too restrictive. Whether the evidence a communicator provides of the first, basic layer of information is direct—as in (6d)—or indirect—a linguistically encoded form as in (6e)—we are still dealing with intentional communication.[8] (7a) below, then, does not exclude (7b), and the distinction between

[8] Sperber & Wilson (1986/1995: 30)—following Schiffer (1972) and Grice (1982/1989)—argue that in another way Grice's communicative intention is not restrictive enough, since it fails to exclude cases in which a communicator provides evidence of an informative intention, and intends an audience to infer this intention, but is not being completely open about their communicative intention. In response to this, Sperber

intentional communication and other forms of information transmission, crosscuts Grice's natural/non-natural distinction:

(7a) Mary's pale complexion means$_N$ she is unwell.
(7b) Mary intentionally communicates she is unwell by deliberately and openly drawing attention to the fact that she is pale.

This raises the question of the extent to which naturally occurring behaviours—natural signs in Grice's sense—may figure in intentional communication. I am talking 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 the next section.

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[9] betray the true feelings of someone we are talking with. Perhaps they cover their face with their hands, or turn away, or hold a book 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

---

& Wilson (1986/1995: 61) propose that the communicative intention is better analysed as an intention to make an informative intention "mutually manifest" to communicator and audience.

[9] 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).

displaying their natural behaviour their informative intention is made evident to both communicator and audience.[10]

It's worth stressing 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[11]; Grice himself was perfectly happy to regard the voluntary production of otherwise natural behaviours (voluntary frowns in the example he discusses (1989: 219)) as meaning *non*-naturally, and he was surely right on this. The deliberate production of an otherwise natural behaviour (in order to make evident an informative intention) is a clear example of intentional communication. It is natural, involuntary, spontaneous behaviours that are the focus here, and *their* status with regard to intentional communication that is being considered.

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. In doing this, Lily draws attention not only to the fact that she is cold, but, by providing evidence of her informative intention, also to the fact that she is cold enough to want to go inside. I think this example is 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 opposed to deliberately shown. In which case, as we have seen, there is no doubt that the shiverer can be said to be communicating intentionally.

---

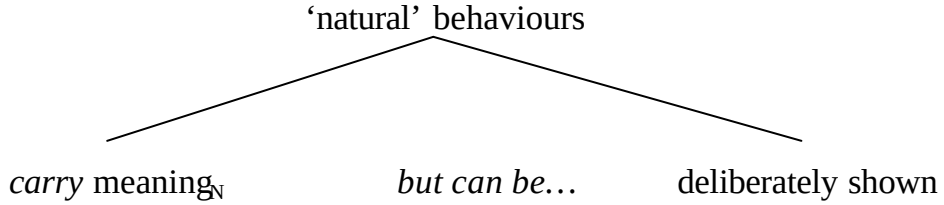
[10] I say '*might* intend' because I would not like to deny that there are probably cases in which crying openly does not amount to an intentionally (deliberately) shown act—expression of grief at a funeral, for example.

[11] 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.

The third example is a smile, which together with facial expressions in general will be the focus of the next section. If we accept that crying 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 is made easier because 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 (Tschiassny 1953, Rinn 1984), and physiological evidence supports the claim first made by Duchenne (1862), and later Darwin (1872), that the muscles used in spontaneous smiles are different to those responsible for deliberate ones. Ekman (1997) 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 mean *non*-naturally, 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 which might, from a Gricean viewpoint, be regarded as simply cases of natural meaning, can also be recruited for use 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 introduction, 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, and are routinely deliberately shown. This also suggests that his natural/non-natural distinction was not exhaustive, and means that the right-hand side of the above diagram will require some fine-tuning.

3 Relevance theory and ostensive-inferential communication

Although he did not talk in terms of a 'coding-decoding model' or 'the code model', Grice's characterisation of meaning$_{NN}$ provided the first alternative to this conception of how humans communicate. This code model view of linguistic communication can be traced back to Aristotle's 'common language'; it is the view taken in the semiotics literature, and one that is still pervasive (see Fodor 1975: 106, Millikan 1984).

Under the intention-based or *inferential* pragmatic models his work inspired, communication is achieved by a speaker giving evidence of an intention to inform the hearer of something, and the hearer inferring this intention. According to relevance theory, language is a code, governed by a modularised, autonomous mental grammar—the language module (Chomsky 1986, 1995). However, utterance interpretation is a two-stage process. The linguistically encoded form, which represents the output of the mental grammar, is simply a starting point for rich inferential processes guided by expectations of relevance.

Relevance theory combines Gricean intention-based pragmatics with aspects of modern psychological research and cognitive science to provide a cognitive-inferential framework. It is built on two principles. *The Cognitive Principle of Relevance* makes a fundamental assumption about human cognition: humans are geared to look out for *relevant* information, which will interact with existing mentally-represented information and bring about cognitive effects in the form of inferences that would not otherwise have been possible. Relevance itself is a property of inputs to inferential processes, and is defined in terms of cognitive effects gained and processing effort expended: the more cognitive effects gained, and the less processing effort expended in gaining those effects, the greater the relevance.

The human disposition to search for relevance is an evolved consequence of the need to process information as efficiently as possible. It is, furthermore, a disposition that is routinely exploited in human communication. Since speakers know that listeners are looking out for relevance, they make their communicative stimuli as relevant as possible (relative to their own goals). *The Communicative Principle of Relevance*

takes it that by making evident an intention to inform—providing an *ostensive* stimulus—a communicator communicates 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.

This communicative principle motivates the following comprehension procedure (taken from Wilson (2000: 420-1)):

> *Relevance theoretic comprehension procedure*: follow a path of least effort in computing cognitive effects.
> (a) Consider interpretations in order of accessibility.
> (b) Stop when your expectation of relevance is satisfied.

Sperber & Wilson propose that the meta-communicative abilities necessary for use of this comprehension procedure might form a sub-part of the meta-psychological ability known as *Theory of Mind* or *mind-reading.* There is thus an interesting point of contact between the philosophical literature (Grice 1957, Strawson 1964, Searle 1969, Schiffer 1972), the intention-based pragmatic models that Grice's work has inspired (such as relevance theory) and more recent psychological research on the capacity to attribute mental states (Leslie 1987, Happe 1994, Baron-Cohen 1995, Scholl & Leslie 1999).

One parallel between relevance theory and the 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.* However, the two are not exactly parallel (see Carston (1998) for further discussion). In relevance theory, *explicatures* are recovered via a mixture of linguistic decoding *and* inference. The basic explicature, the proposition expressed (roughly equivalent to Grice's *what is said*), is rarely recovered—as Grice seems to have thought—by disambiguation and reference assignment alone, and the construction of *higher-level explicatures* requires even more pragmatic development, such as the embedding of the basic truth-conditional content under a speech-act or propositional-attitude description.

To illustrate this approach, consider how Mary's mother might interpret Mary's utterance in (8a). Having recovered the basic explicature—the proposition expressed—she might embed it under a speech-act description, as in (8b). Or, according to the paralinguistic clues provided by Mary's tone of voice, she might embed it under a propositional-attitude description, as in (8c). These would be higher-level explicatures of Mary's utterance.

(8a) Mary (in a regretful tone of voice): I don't feel well.
(8b) Mary is saying that she doesn't feel well.
(8c) Mary regrets that she doesn't feel well.

As has already been pointed out, in contrast to Grice's primarily semantic concerns, relevance theory aims at providing a characterisation of human intentional communication generally. Utterances, after all, are not the only kind of ostensive stimuli, and a communicator might choose to draw her audience's attention to her intention to inform by means of a look, a gesture, even a natural sign. Relevance theory, therefore, does not attempt to draw the line that Grice wanted to between "'deliberately and openly letting someone know' and 'telling'" (1989: 218). Instead, it recognises that both *showing* and *saying* are instances of ***ostensive-inferential*** communication. This fits with the account of 'intentional' communication presented in Section 2, and from now on I will adopt Sperber & Wilson's terminology.

Recall our earlier characterisation: in any act carried out with the intention of providing evidence of an *informative* intention there are two layers 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 communicative act a case of 'showing' or 'saying' is the nature of the evidence provided for the first layer: in cases of showing, the evidence provided is relatively direct—a natural sign for example; in cases of saying, the evidence provided is indirect—a linguistic utterance. And rather than the dichotomy Grice envisaged, there is a continuum of cases between the two extremes. Firstly, most cases of showing—cases in which the evidence provided is fairly direct—still require an extra layer of inference before the communicator's full informative intention is recognised (recall that in (6d) Mary's mother still has to infer that Mary wants help), and the extent to which an audience must rely on this extra inference is a question of degree. Secondly, the ostensive stimulus is usually a composite of different, inter-related behaviours anyway: a *combination* of showing and saying.[12]

The ostensive stimulus, then, is made up of those behaviours that the audience recognises as deliberately shown. In (8a) although Mary's regretful tone of voice (and the slight frown that may accompany it) are 'natural', they are deliberately shown to provide evidence of her informative intention and information about her attitude to the

[12] In more ways than one, since, being ostensive acts, cases of saying actually *are* cases of showing. I abstract away from this issue here (it is discussed in Wharton (forthcoming)).

proposition she is expressing, or about emotions or feelings. To answer question (b)—*what do these natural behaviours indicate?*—these ostensive paralinguistic behaviours, when accompanying linguistic acts, contribute to the construction of higher-level explicatures. Of course, Mary's utterance might be accompanied by other natural signs: perhaps she is pale, but doesn't realise; perhaps she is covered in spots. These other signs (presuming they are noticed) will still be processed by Mary's mother, 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 outlined above. Sometimes, of course, natural behaviours are used alone as ostensive stimuli, and do not accompany utterances. Since I will argue that the question of whether, in such cases, they communicate at an explicit or implicit level turns largely on *how* these behaviours communicate, I leave discussion of this until later.

In the next section, I turn to questions (c) and (d). I will argue that natural behaviours do not all work in the same way, and that some are inherently communicative. These, I propose, are best analysed as *natural codes*.

4 Natural codes

4.1 What is a code?

Broadly speaking, there appear to be two different answers to this question. 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. The semiotician Thomas Sebeok, summarising Shannon & Weaver's classic (1949) model, puts it like this:

> 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: 12-13 (my emphasis, TW)).

On the other side, there is the notion of code most often used in the social sciences, and perhaps more common in ordinary language use. 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.

The two notions do not have to be seen as mutually exclusive. Both, for example, are recruited in explanations of human language. In the first sense, language *is* a system that pairs signals (words) with messages (meanings). In the second, many people (including Lewis, and Grice himself) see language as a set of signalling conventions. Grice, for example, saw his natural/non-natural distinction as a reformulation of that between natural and 'conventional' signs (1989: 215).

However, I use the word 'code' is the first sense: the strict semiotic sense. Furthermore, I use it to the exclusion of the second sense. Given the Chomskyan perspective on language adopted in this paper, 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[13] 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: 84).

While many will disagree with an attempt to sever the link between language and convention, few would advocate talk of conventional codes when discussing animal communication systems. However, in the first, strict semiotic sense, codes they most certainly are. In that regard, the existence of *natural* codes is uncontroversial.

4.2 Bee-dances and smiles

Consider honeybees: the honeybee performs a complex dance in order to indicate to its conspecifics information relating to the distance, location and quality of nectar (von Frisch 1967). Would we want to characterise the 'meaning' the dances have as meaning_{N} or meaning_{NN}? Recall the tests: 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

[13] Quoted in Davies 1996.

bee has performed the dance means that the nectar is there[14]; 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 bees, then, mean naturally.[15]

There is, nonetheless, a sense in which it is plainly unsatisfactory to see the meaning of the bees' 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 bee dances 'mean' something to other bees: 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 bee-dances. The results, which in (4ab) and (5ab) illustrate so neatly the "reasonably clear intuitive distinction" he sought to demonstrate, become strangely unreliable (9ab):

(9a) What is meant by the bee's dance is that there is nectar at location$_X$.
(9b) That dance means "nectar-at-location$_X$".

At least in the case of honey-bee dancing, it appears that forms of paraphrasing previously only appropriate to describe cases of meaning$_{NN}$, are also appropriate in cases that other tests suggest are cases of meaning$_N$. Why the tests become unreliable is unclear, but it seems to me that the most plausible explanation is that our intuitions concerning (9ab) simply reflect the fact that the dances of bees are inherently communicative: they are *coded* signals. Bees don't 'mean' as in 'intend' (see *fn.* 2),

---

[14] This is not to under-estimate the complexity of bee signalling, which 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 & Bekoff 1997: 108).

[15] I had originally wanted to include another example from the world of non-human animal communication at this point: the alarm calls of vervet monkeys, for example. However, it has been claimed (Seyfarth et al. 1980) that vervets do not automatically emit an alarm call on seeing a predator: this suggests it is not factive (although it would still be factive if they *only* emitted the alarm call when there was a predator). Furthermore, it has been suggested that vervets have some sort of rudimentary meta-representational abilities, which suggests they may be able to attribute 'lower'- order intentions. I don't believe that the existence of these intermediate cases undermines my argument. On the contrary, it confirms my suspicions that a strict natural/non-natural distinction, as proposed by Grice, is not exhaustive.

but something is surely meant (in one sense) by their dances (hence von Frisch's pioneering work). In this sense, then, Grice's dichotomy is not exhaustive.

Now Grice's 1957 paper was not remotely concerned with non-human animal communication, and it could be argued 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 (9ab) raise have implications beyond the rather trivial observation that Grice's natural/non-natural dichotomy fails to accommodate the dances of bees. 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 2: (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 (see (3ab)): 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 (10abc) below:

(10a) The fact that he is smiling means he is happy.
(10b) The fact that he is crying means he is unhappy or distressed.
(10c) 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[16] 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—used to mean non-naturally. 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

[16] 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.

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 bee-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 (11ab) are acceptable in a way that (11cd) possibly, and (11ef) certainly, are not.

(11a) His smile means "I am happy"
(11b) What was meant by that smile was that he is happy.
(11c) ?His tears mean "I am distressed or unhappy'.
(11d) ?What was meant by his tears was that he is distressed or unhappy.
(11e) *His shiver means "I am cold".
(11f) *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 (12ab) *would* be appropriate paraphrases:

(12a) Her (sardonic) smile means "I am not amused".
(12b) What was meant by her (sardonic) smile was that she was not amused.

And it seems clear to me that (11ab) are acceptable in a manner distinct from this, a manner more closely akin to the cases in (9ab).

Just as the dances of bees '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 bee dancing, the reason for this, I claim, lies in the adaptive functions of the behaviours themselves.

4.3 Signs and signals

In order to clarify which instances of information transmission in the animal world are to be regarded as communication, and which are not, Hauser (1996: 9-10)

distinguishes between two ethological notions: *signals* and *signs*.[17] *Signals* are those behaviours that have been designed, in the evolutionary sense, to convey information. The dances of bees 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.

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, 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 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[18, 19]: natural behaviours, then, do not all work in the same way.

---

[17]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.

[18] 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 (11a-f) as applied to crying as I am as applied to shivering and smiling; secondly, the dissociation between crying and tears makes it hard to say

In the ethological literature, non-human animal communication systems are often referred to as codes (see Bradbury & Vehrencamp 1998: 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*. 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 all-purpose 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.

In fact, although Shannon & 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 bees' 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

---

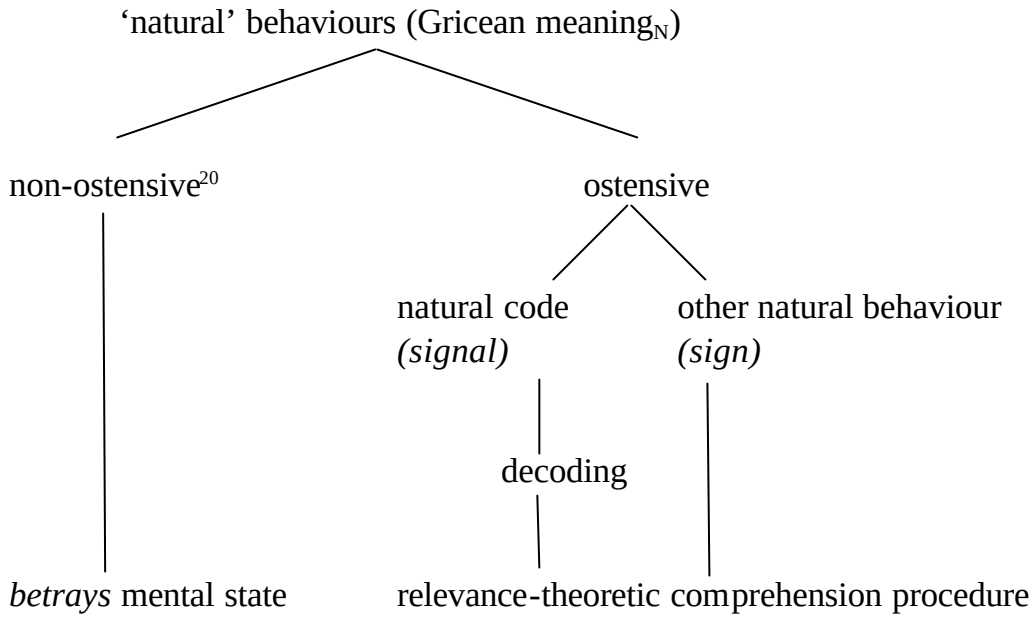
with any authority exactly what the adaptive value of crying (as distinct from shedding tears—on which see Darwin (1872-1998: 164-175)) is.

[19] 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*. 2 overleaf—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). In fact, I'm not sure how Clark would deal with ostensively-used symptoms, given that the notions he regards as central to communication are 'common ground', 'joint purpose' and 'coordination', rather than the recognition of intentions.

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.

For natural coded behaviours are routinely put to use in ostensive-inferential communication. When this is the case, 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 2 with the observations made here, we can represent the new picture as in *fig.* 2 below:

fig. 2



The work of Paul Ekman (1989, 1992, 1994, 1999) suggests 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: 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).

---

[20] To complete the diagram there should perhaps be a version of the signal/sign distinction under this 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 pragmatics to be intentional—ostensive—communicative acts. The focus of this paper, 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 infer from their putatively non-ostensive behaviour that they are happy, there is still the question of why they are smiling in the first place. As I have tried to show, it is not enough to offer 'people smile when they are happy' as an account of why people smile in the way that we might offer 'people shiver when they are cold' is an account of why people shiver. In evolutionary terms, the function of smiling is to *display*: they are (or have become), we might say, *ostensive-by-design.*

Fridlund's *behavioural ecological* view, for example, stresses the manipulative communicative function of facial expression, whilst denying they are reliably correlated with the expression of 'emotion'. Among the evidence he presents in support of this position are data from experiments on 'audience-effects' in human smiling (Kraut & 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: 153).

Hauser (1996: 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 emotion in the production of facial expression, 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.

Exactly what kind of information smiles and spontaneous expressions of emotion might encode is unclear: "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 person's face" (Ekman 1989: 159). In the next two sub-sections, I examine previous attempts to clarify this.

4.4 Sebeok's *analog* codes and Peircean *indices*

The approach to human interaction within the *semiotic* program (Peirce 1897, 1903, de Saussure 1916/1974) was that of a strict coding-decoding model. It proposed that not only linguistic communication, but customs, rites etc., were best analysed as systems of signs, or codes. Sebeok (1972) investigates the different types of coding humans might use in linguistic communication, and proposes that there are two kinds: *digital* and *analog*. His distinction was intended to reflect what he termed 'rational' and 'emotional' (1972: 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 analog-digital distinction exists in a variety of guises.[21] 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 analog 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.

An analog 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, 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.[22] 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 a quivering needle, the engineer consults a numerical, i.e. digital read-out.

The principal advantage of digital over analog is one of accuracy: digitally encoded information is 'all-or-none'; analog information is 'more-or-less'. 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'[23], and humans are capable of producing (and understanding) an unlimited number of distinct combinations, expressing an infinite range of properties. Within an analog system there are no

---

[21] Hauser (1996: 54) credits the cognitive ethologist Peter Marler with being the first to propose the distinction as applied to communicative systems.

[22] The pressure gauge analogy is a development of one presented in Bolinger (1983).

[23]To paraphrase Chomsky's (1988: 169) famous terminology.

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.

The potential that analog and digital codes have to complement each other is nowhere more clearly illustrated than by human linguistic communication. Consider (13ab) and (14ab) below:

(13a) Lily: Has John arrived?
(13b) Jack: John has arrived.
(14a) Lily: Has John arrived?
(14b) Jack: (smiling happily, in a pleased tone of voice) John has arrived.

By fronting the auxiliary in (13a), Lily forms an interrogative. One aspect of the difference in intended meaning between her utterance in (13a) and Jack's in (13b) is indicated by placing the discrete units of language in a different order. But consider Jack's replies in (13b) and (14b). 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 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 analog 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.

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 uttered in an entirely neutral tone of voice would sound extremely strange, and it is unlikely they would communicate his feelings very effectively. In everyday communication we simply take for granted how we 'naturally' reflect our emotional intensity.

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—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" (2/299).[24] An indexical representation is one in which the relationship is physically (or temporally) proportional or causal—the analog pressure gauge described in the previous section is a good example: "it is physically connected with its object…they make an organic pair" (2/299). 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" (2/299).

Sebeok attempts to draw parallels between Peirce's distinctions and the analog-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 structuring in the code" (1972: 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[25], we might look at the extent to which analog 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 analog 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 an encoded representation of a dog, simply a likeness, but the causal and temporal links of Peircean indices reflect the kind of proportional relationship mentioned earlier as characteristic of analog coding. This observation is reflected in a number of previous approaches. Jakobson & Halle (1956: 11) describe the expressive features of speech as "physiognomic indices". In an analysis with strong echoes of Peircean indices[26],

---

[24] Quotes from Feibelman's *'An Introduction to Peirce's Philosophy'* (pp. 91-92). References to Peirce's *'Collected Writings' Volume 1-6*.

[25] Are there, for example, arbitrary *analog* codes?

[26] Although it is entitled 'The inherent *iconicism* of intonation' (my italics, TW).

Bolinger (1983) describes intonation generally as exhibiting properties of "dynamic indicators".[27]

The concepts of analog codes and indices are certainly useful notions. The majority of human natural codes do indeed appear to work along analog lines, and the notion does at least take us some way towards an answer to question (c)—*how are they interpreted?* However, if we are to attempt an answer to question (d)—*in cognitive terms, what kind of information do 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 analog codes and indices.

Firstly, semiotic accounts rely entirely on a code model of communication; but a coding-decoding model 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 bees. Furthermore, semiotic approaches aspire to providing a 'general' theory of communication (albeit in terms of coding and decoding). As we have seen, however, 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 analog is too restrictive. Codes of the kind used by vervets (Cheney & Seyfarth 1990) and 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,

---

[27] Bolinger's analysis of intonation might also be seen as an attempt to characterise it as one kind of natural code: "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: 98). On the question of sifting the linguistic from the non-linguistic, which I mentioned in my introduction, he writes— "Intonation… assists grammar—in some instances may be indispensable to it—but it is not ultimately grammatical" (1983: 106), and concludes, "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: 108).

since 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 analog 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, in cognitive terms, what the information encoded in a human natural code looks like.

4.5 Wierzbicka: The 'semantics' of facial expressions

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: 178); "the basis for decoding lies either in similarity…or in co-occurrence" (2000: 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.[28]

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 above in Section 3. 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 sense that the time has come for a change of paradigm" (2000: 147). She lists ten basic assumptions fundamental to this new direction.[29] These assumptions lay the foundations for what is essentially her

---

[28] Though this is not to suggest they might not somehow *become* coded.

[29] 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: 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*

main claim, that the coded element of human facial expressions can be analysed in terms of her own 'Natural Semantic Metalanguage' (NSM).

The NSM approach is based on a set of around sixty primitives, which it is claimed represent the innate components of meaning: "All complex meanings, in all conceptual domains, can be represented and explained as configurations of these sixty or so fundamental conceptual building blocks" (2000: 153), among these—'I', 'you', 'someone', 'want', 'think', 'feel', 'do', 'good', 'bad', 'small', 'place'. These fundamental building blocks are part of the human genetic endowment, and are therefore universal as well as innate. Wierzbicka and her colleagues use the approach to account for a variety of 'semantic' issues: "the indefinable concepts—the primitives—are the fundament on which the semantic system of a language is based" (1996: 14). Below is an example of the kind of complex conceptual structure that NSM analyses propose for words (15) (Wierzbicka 1996: 241):

(15) *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

---

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 & 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: 159). I think Hauser's point about two different levels of analysis (see p. 23 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.

Wierzbicka has also extended the approach to the meaning encoded by interjections. See (16) below (1992:168):

(16) *yuk*
I now know/imagine something
I think: it is bad
I feel something bad because of that
I think other people would feel the same
I feel like someone who thinks
I don't want to be in the same place as this

And it is now applied to facial expressions—(17) below:
(17) *raising of the eyebrows*
I know something now
I want to know more (about this)
I'm thinking now

In Wharton (forthcoming) I argue against NSM-based accounts of meaning generally. The approach, for example, is massively decompositionalist, and I am largely convinced by the philosophical and psycholinguistic arguments presented by Jerry Fodor and his colleagues (Fodor, Fodor & Garrett 1975, Fodor 1981) that such accounts of word meaning are highly problematic. These arguments apply equally to NSM-based accounts of interjections, which are the focus of that paper, and extend to any approach which suggests a decompositionalist approach to the meaning conveyed by facial expressions. I can, for example, see no reason why *raising the eyebrows* should encode 'I'm thinking now'. There are so many exceptions—most obviously, that (in my experience at least) people actually tend to furrow their brows when they are thinking—as to render an account based on the structure in (17) unworkable. Decompositionalism doesn't work for facial expressions any more than it works for words.

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 made in the introduction, and is a similar criticism to the one levelled in the last section against the semiotic program generally.

The NSM approach to human communication is highly reminiscent of just the kind of coding-decoding model which I am rejecting. Wierzbicka (1996: 8) talks in terms of "the meanings encoded by natural language", and largely ignores the kind of insights discussed in Sections 2 and 3 concerning the link between intention and meaning.

Consider the following on facial expressions: "Human faces send messages, and these messages *must be decodable*" (2000: 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 (perhaps along the lines of those sketched in Section 3), 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 NSM 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: 151). There is a semantics/pragmatics distinction in NSM 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 (though see Ekman 1999) that has been much discussed. Eibl-Eibesfeldt (1972: 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: 300). Wierzbicka comments that her structure is not inconsistent with interpretations relying on the technical expression 'attentional activity' (2000: 168).[30]

But it is hard to see how. The gap between the conceptual structure in (17) 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.[31] NSM 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—analog/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: 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 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').

---

[30] Though she does not endorse the use of the expression. Wierzbicka is actually referring to a paragraph from Smith & Scott (1997: 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: 164).

[31] Allowing for my previous reservations over iconic representations being coded.

This element of confusion is unavoidable within an NSM 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 which 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: 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: 270).[32]

However, what these criticisms also reveal is a Whorfian, deterministic 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: 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.

Sperber & Wilson (1998) argue 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

---

[32] Wierzbicka's response to this quote from Ekman is to quote him again, this time from his rebuttal of Russell's critique (1994: 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: 273-275) for his response.)

mind if there are public words with which the information matches. They argue that, in fact, "there are many times more concepts in our minds than words in our language" (1998: 198), and that in everyday thought we regularly entertain unlexicalised concepts. We communicate them regularly too, for a concept that is lexicalised may be inferentially enriched by a hearer to derive the (slightly different) precise sense that a speaker intended to communicate. To the extent that it is correct that encoded concepts may be enriched in this pragmatic-inferential manner, 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.

As it stands, I think these three arguments suggest that NSM 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? I believe it does, and should be explored. Ultimately, I believe it represents the final nail in the coffin of the NSM account: the information encoded in natural codes is not *conceptual* at all.

4.6 Translational and non-translational activation of concepts

Over the past 30 years, linguists and philosophers have explored the idea that not all linguistic meaning is of the same type. Speech-act theorists (Austin 1962, Searle 1969, 1979, Bach & Harnish 1979), for example, proposed that sentences both express propositions, which *describe* the world, and may contain non-truth-conditional expressions, which *indicate* the speech (illocutionary) act a speaker is intending to perform, or the propositional attitude a speaker is expressing. Although the sentences in (18abc) all express the same proposition—*Mary goes to the doctor*, they differ in their illocutionary force: (18a) may have the force of a question; (18b) of a request for action; (18c) of an assertion.

(18a) Does Mary go to the doctor?
(18b) Mary, go to the doctor!

(18c) Mary goes to the doctor.

Speech-act semanticists claimed that their approach could deal with a whole range of non-truth-conditional linguistic expressions, including mood indicators—the linguistic features which mark interrogatives and imperatives (word order in the above examples)—and various adverbials which do not contribute to the truth-conditional content of an utterance, for example those in (19ab):

(19a) *Regrettably*, I don't feel well.
(19b) *Happily*, John has arrived.

Distinctions that reflect aspects of the describing/indicating dichotomy have been proposed in a wide range of linguistic analyses. Grice's own analysis of discourse connectives such as 'but' and 'moreover' owed a great deal to the speech-act distinction between describing and indicating. He proposed that 'but' conventionally implicates the performance of higher-order illocutionary speech acts. So while a speaker might be asserting (20b) and (20c) in an utterance of (20a), what she is conventionally implicating is that the two assertions are to be contrasted.

(20a) Mary is unwell but she's going to school.
(20b) Mary is unwell.
(20c) Mary is going to school

Discourse pragmatic analyses of connectives propose that they "signal relations between units of talk" (Schiffrin 1987: 40) rather than contribute directly to the units of talk themselves.

In formal semantics, David Kaplan (1989) has suggested that the pronoun 'I' encodes a rule of use**,** which he calls its *character*, which can be distinguished from the propositional (descriptive) *content* the pronoun contributes; so while the content of a given utterance of 'I' will be the person it refers to, the character is the rule for identifying the referent in any given context. Kaplan also distinguishes the (non-truth-conditional) *expressive* content of an utterance from its descriptive content. Typically, what expressives display is something about the state or attitude of the speaker. This distinction closely mirrors speech-act distinctions—Searle (1979), for example, would analyse the adverbials in (19ab) as performing expressive speech-acts.

The question of what exactly these various distinctions amount to in cognitive terms is one that any cognitive pragmatic theory needs to answer. Code theorists have few

options. Since the only way to communicate a concept is to encode it, words with descriptive content and words that indicate must convey the information they do by encoding it conceptually. In an inferential model, however, there is another possibility. It is this possibility I explore in this section, and I will argue that the analysis I outline carries over to the natural coded behaviours that are the focus of this paper.

Broadly speaking, what a communicator *C* wants to communicate to her audience *A* is a thought. I take it, following Fodor (1983) and others, that thought takes place in some sort of modality-neutral representational medium. This neutral medium takes the form of structured strings 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. If humans were truly telepathic, *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 behave ostensively in such a way as to get *A* to entertain the same (or, more accurately, a similar[33]) 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, thoughts are communicated along the lines presented in the Sebeok quote above: *C*'s thoughts are translated into the code, and translated back from the code by *A* (see page 14-15). As I argued earlier, this kind of account is fine in the case of bees, but it does not adequately characterise human linguistic communication.

That being said, there is of course a coded element to human linguistic communication: some words *do* translate into the constituents of thoughts, or concepts. If I utter the words 'bear' or 'hit' or 'near' 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 'urso', 'bater' or 'perto'.[34] I call this the ***translational*** activation of concepts, and the kind of coding that gives rise to it I call translational coding.

---

[33] One effect of a shift to an inferential model is that the thought A entertains as a result of communicating with C will only *resemble* C's original thought, rather than duplicate it. Paradoxically, a code model view of communication results in a kind of thought-transfer similar to the kind one would expect if humans were indeed telepathic. Sperber (2001) calls it "cognition by proxy".

[34] Another effect of a shift to an inferential model is that it suggests an inferential solution to the problem of context-dependency generally: why is it that in 'the bear walked out over the ice', 'bear' would usually be taken as communicating 'polar bear'?; how can the verb 'hit' have such clearly different senses in 'I hit my boss' and 'I hit my head'?; why is it that if I'm talking to an Australian I live 'near' London

However, as speech-act theorists recognised, linguistic meaning is not of just one type. In cognitive terms, this raises the question of whether all words work in the same way. I propose they do not, and 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. I call the kind of coding that gives rise to this kind of activation non-translational coding, and will propose that this is the kind of coding natural codes exploit.

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.

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 leave it entirely implicit—to be interpreted inferentially—and not encode anything. Recall the examples from the last section where I deliberately and openly show you my shiver, intending to communicate that I feel cold; or when 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 (20):

(20) Jack: (referring to John) He's arrived.

Jack has clearly encoded something by uttering the pronoun 'he', but what? Wilson & Sperber (1993) argue that, from a cognitive point of view, Kaplan's analysis of

---

(50 miles), but when talking to a Londoner—although I still live the same distance away—I don't? I discuss the inferential solution briefly (inferential enrichment or loosening of encoded concepts) on p. 34 above. On the issue of context-dependency, and the implications of adopting a fully inferential approach, see Wilson 1998.

pronouns in terms of character and content amounts to the claim that rather than encoding a conceptual representation of the intended referent[35], what pronouns actually encode is a *constraint* which helps the hearer identify the intended referent in a given context by making it 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 utterances we intend—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 inferentially. Linguistic indicators are coded signals, but the code is non-translational: their function is to guide and constrain inference.

One way to cash out the distinction between translational and non-translational coding has been explored in relevance-theoretic semantics in terms of a distinction between words that encode concepts, and those that encode *procedures*. The distinction was first proposed in Blakemore (1987).[36] Most words encode 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 the constituents of conceptual representations, the function of these words in Blakemore's view is to constrain the inferential processes involved in constructing or manipulating these representations during the search for relevance. They guide the comprehension process by narrowing the hearer's search

---

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

[36] Blakemore (1987) doesn't discuss the conceptual/procedural distinction in terms of one between translational and non-translational activation of concepts. Nonetheless, it seems appropriate to draw a comparison between the two distinctions: firstly, linguistic expressions such as pronouns, mood indicators, particles etc., which it has subsequently been proposed encode procedural information (see Wilson & Sperber 1993), *clearly* result in the non-translational activation of concepts along the lines I am suggesting; secondly, while I would agree that discourse connectives perhaps don't provide the best examples of non-translational activation (that's why I chose to present the idea using pronouns), that is clearly how they function: if they don't map onto concepts, but are responsible for the *construction* of conceptual representations, then it's hard to see how the activation of the concepts *in those representations* is anything but non-translational.

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 analyses of the discourse connectives, 'so' and 'after all'. Two possible interpretations of (21a) would be spelled out more explicitly in (21b) or (23c):

(21a) Mary is pale. She is unwell.
(21b) Mary is pale; *so* she is unwell.
(21c) Mary is pale; *after all* she is unwell.

On Blakemore's account, in (21b) 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 (21c) the expression 'after all' encodes a procedure which leads to the second proposition being understood as evidence for the first.

Wilson & Sperber (1993) develop these proposals, and as well as suggesting a procedural account for pronouns, they propose that mood indicators might be analysed as encoding procedural information which indicates the attitude of the speaker to the thought she is conveying.[37] They also propose that certain attitudinal discourse particles and interjections are best analysed as encoding procedural information.

In Wharton (forthcoming) I use the translational/non-translational distinction as a way of fleshing out the conceptual-procedural distinction. However, it is possible to see things the other way round, and this is the view I intend to explore here. Seen in this way, the conceptual-procedural distinction may be viewed as just one way of construing the translational/non-translational distinction. In Wharton (forthcoming) I explore the possibility that interjections—items with only marginal linguistic status, such as *wow*, *aha*, *oh*, *yuk*—might be analysed in non-translational terms. Interjections, like mood indicators, express attitudes (though not always propositional ones). On my account, 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. Utterances of (22a) and

[37]See Wilson & Sperber (1988) for arguments against the speech-act account of mood indicators and a presentation of their account—though without reference to procedural encoding

(22c) might lead the hearer to form the higher-level explicatures in (22b) and (22d) below:

(22a) *Aha*! John's arrived.
(22b) The speaker is surprised that John has arrived.
(22c) *Wow*! John's arrived.
(22d) The speaker is delighted that John's arrived.

The utterances in (22a) and (22c) might be compared with their conceptually-encoded counterparts in (23ab):

(23a) I am surprised that John has arrived.
(23b) I am delighted that John has arrived.

Interjections, of course, are fairly exceptional in that they can constitute an utterance in their own right, in a unique non-elliptical manner. In this case, rather than there being any propositional embedding—there is, after all, no propositional content to embed—an interjection simply encodes a fairly vague pointer to the emotional/mental state of the speaker.

Apart from the intuition that they indicate rather than describe, what theoretical evidence is there to suggest that interjections do not encode this information conceptually? Firstly, in common with some other expressions that encode non-translational information, interjections do not contribute constituents directly to the truth-evaluable content of the utterances that contain them.[38] One way of explaining this is to see it as a reflection of the fact that they provide information on how to process conceptual representations, as opposed to actually entering into the representations themselves. Consider (24ab) below:

(24a) I'm delighted. John has arrived.

---

[38] This is not to suggest that all non-truth-conditional meaning is procedural (see Ifantidou-Trouki 1993). In fact, one of the interesting consequences of this approach is that it suggests that it is not the case that all non-truth conditional indicators—such as mood indicators, adverbials and connectives—work in the same way. Consider, for example, the 'indicators' in (20ab), which appear to encode conceptual information. This entails an important modification to speech-act analyses, for the conceptual/procedural—or, for that matter, the translational/non-translational—distinction crosscuts the describing/indicating distinction.

(24b) X IS DELIGHTED - JOHN$_X$ HAS ARRIVED.[39]

An utterance of (24a) makes two assertions. Intuitions tell us that it is true when and only when the speaker is delighted and John has arrived. These intuitions reflect the contribution of the words in (24a) that *do* encode concepts to the conceptual representation a hearer of (24a) might be led to entertain—(24b). By contrast, (22c) only makes a single assertion; intuitions tell us it is true if and only if John has arrived. Whilst a hearer of (24a) might respond with an utterance of (25) below, for example, it seems implausible to suggest that it might be uttered in response to (22c).

(25) That's not true; you're not delighted.

This is not to deny that a hearer of (22c) might be led to entertain a conceptual representation such as (24b), or the higher-level explicature in (22b); nor is it to suggest that "the meaning conveyed" (to use Wierzbicka's term) by an utterance of *wow* is not to some extent paraphrasable by a phrase such as 'I am delighted'. It is just to suggest that the route to the appropriate representation is different (i.e. non-translational). This argument applies equally to other interjections, and suggests that it would be surprising if they encoded concepts.

Secondly, interjections are not linguistically productive: they do not compose to form larger phrases. If words encode concepts, it is important that they do, in fact they *must* in order to translate into the language of thought and form the structured strings that are complex conceptual representations.[40] In the case of words encoding non-translational information, there is no such expectation, for their function is not to translate into the language of thought, but rather to point in the direction of appropriate conceptual representations[41]: in the case of pronouns, by narrowing down the set of possible referents; in the case of connectives, by providing information on how to manipulate conceptual representations; in the case of mood indicators, particles and interjections, by providing attitudinal information.

---

[39] I follow convention here, and use upper case for conceptual representations.

[40] This is one reason to think that the adverbials in (20ab) do encode concepts: *regrettable-regret-regrettably*; *happy-happiness-happily*.

[41] In relevance theory terms, "the direction in which the relevance of the utterance is to be sought" (Sperber & Wilson 1986/1995: 254).

There are other reasons to prefer this kind of account of interjections to a fully conceptual one (see Wierzbicka 1992, Wilkins 1992 and Ameka 1992 for a fully conceptual account of interjections). What interjections such as *wow*, *aha*, *oh* communicate is typically quite vague and context-dependent.[42] The vagueness is partially explained if, rather than encode a concept, what an interjection encodes is information that non-translationally activates in the hearer a mental state—a range of attitudinal concepts or propositional-attitude descriptions—from which he must choose during the interpretive process. The context-dependence is also partially explained. Rather than simply being the result of decoding, the precise conceptual structure finally arrived at by the hearer is the outcome of several (potentially overlapping) inferential processes that are set in motion, and constrained by the encoded information.

Whether this non-translational account of interjections amounts to a *procedural* one remains to be seen. Suffice to say, if that were the proposal, then the domain of procedural meaning would be considerably extended.[43] Thus far, procedural analyses have only been proposed for properly linguistic items. As I remarked earlier, the linguistic status of interjections is marginal at best, and there is considerable doubt as to whether interjections are part of language proper (see Sapir 1970, Goffman 1981, Quirk, Greenbaum *et al*. 1985, Ameka 1992). If interjections are not part of language—and I believe there are good reasons to think that they are not (see Wharton (forthcoming))—then it suggests a degree of dissociation between non-translational and linguistic encoding (which may or may not amount to a dissociation between procedural and linguistic encoding). In Wharton (forthcoming) I propose that interjections can be seen as encoding non-translational information *whether or not* they are part of language.

---

[42] Of course, words that encode concepts can be vague and context dependent too—'near', for example. Notice, however, that 'near' does contribute to the truth-conditional content of an utterance, and is linguistically productive

[43] In fact, the notion of procedural information itself would be a broader one than that originally proposed in Blakemore (1987), where procedural information is characterised as *instructions* to the hearer. The notion of procedural information as (potentially) vague indicator (or *anything* encoded non-translationally) adopted here is in sharp contrast with this view. The broader view may well be worth exploring, and it could be argued that it is probably required anyway in order to accommodate the full spectrum of linguistic devices currently analysed as encoding procedural information, but that is a theory-internal question which I do not intend to debate here.

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 (8a) and (14b), repeated below as (26ab):

(26a) Mary: (in a regretful tone of voice) I don't feel well.
(26b) 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 (27ab) below:

(27a) MARY REGRETS [THAT SHE DOESN'T FEEL WELL]
(27b) 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 ((26b), for example, is almost entirely parallel to (22a)). The proposal, then, is that the coded element in all manner of para- and extra-linguistic behaviours—Mary's tone of voice in (26a), Jack's smile in (26b), 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 most words 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 (parallel to the earlier example) 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.

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. Goffman (1981) proposes

that sounds exist on a continuum such as this, from entirely natural sounds through interjections to language proper (see Wharton (forthcoming) for discussion).

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 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 bee calculating the distance between the hive and the nectar according to the intensity of the communicating bee's dance.

The question of the relationship between the coding-decoding of natural codes, those meta-communicative abilities that govern the search for relevance in verbal comprehension and those responsible for the wider meta-psychological mindreading ability certainly needs to be addressed. One clue, provided by evidence from those with impaired mindreading abilities, suggests that the former might dissociate from the latter two. As well as experiencing problems in verbal comprehension, autistic subjects, for example, also have problems 'reading' the emotional states of others. Interestingly, however, experimental evidence suggests they perform better reading 'simple'—situation-based—emotions (such as *happiness* and *sadness*) than 'cognitive'—belief-based—emotions (such as *surprise*) (Baron-Cohen, Spitz & Cross 1993). To the extent that these findings (and the suppositions on which they are based) are correct, this might be taken to suggest that while the meta-communicative and meta-psychological abilities in these subjects are impaired, the natural coding-decoding mechanisms remain intact.[44]

[44] Here there would be an interesting parallel with language. Some autistic subjects show relatively normal linguistic development. Some, despite their autism, show a remarkable affinity for language acquisition (Smith & Tsimpli 1995).

5 Conclusion

I posed four questions in my introduction, and would like to summarise briefly my conclusions. I will then close with a few more thoughts that have emerged as I have been writing this paper.

(a) What is the relationship between natural behaviours and intentional communication?

I have sketched a model of communication in which inference plays a central role. Inferential communication is achieved by a communicator giving evidence of an intention to inform, and an audience inferring this intention. The evidence provided might be an utterance, a gesture or some other behaviour. Natural, involuntary, spontaneous behaviours can be used in intentional—*ostensive-inferential*—communication, and the distinction between ostensive-inferential communication and other forms of information transmission cross-cuts Grice's famous distinction between meaning$_{N}$ and meaning$_{NN}$. I am not suggesting that natural behaviours are only ever interpreted as a result of their being used in ostensive communication. Often, for example, natural behaviours can be used as clues to infer the mental/emotional state of an individual who has no intention of communicating at all. The fact that there is this crosscutting, however, reveals an under-explored area of pragmatics, which should be investigated.

(b) What do these natural behaviours indicate?

An ostensive stimulus is typically a composite of paralinguistic as well as linguistic signals. When accompanying linguistic acts, these ostensive paralinguistic 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 Wharton (forthcoming) I propose that in such cases interjections—as is the case with non-verbal communication generally—communicate implicitly. I think there is something inherently problematic in this, and am unhappy to carry this over to my analysis of natural codes (this affects my claims regarding interjections). 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 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. This issue requires further thought if paralinguistic behaviours are to be fully integrated into the relevance-theoretic framework.[45]

(c) How are they 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 & Bekoff (1997) for discussion). 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 & Smiley 1991).

The term 'paralinguistic' 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,

[45] 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.

or recognised as ostensive, it will not be picked up by the relevance based comprehension procedure.[46]

(d) In cognitive terms, what kind of information do natural codes encode?

The information encoded in these natural codes is non-translational in the sense described in the last sub-section. 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 clear 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.

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 aspects of intonation. Non-translational information works in a variety of different ways which reflect the differences between the linguistic devices and non-linguistic behaviours that encode it.

*

A recurring theme of this paper has been to argue for the need to see human communication in terms of an inferential model. We underestimate the intentional aspects of human communication at our peril. One general issue that has come out of my research for this paper is that the notion of intentionality ("rich, philosophical sense"—see *fn.* 2) is largely (and conspicuously) absent from the literature on human facial expression. 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" (Fridlund 1994: 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

[46] Assuming it can make such subtle distinctions. There might, for example, be a situation where the procedure is triggered 'accidentally', as it were.

Marceau of expressiveness, modern humans" (Hauser 1996: 265). But what distinguishes human from non-human animal communication is not that we live in a world populated by other faces, but that we live in a world populated by other minds.

There are other areas on which a framework such as the one presented here might shed some light. Returning to the continua proposed by Kendon and Goffman, and briefly discussed in the last section, these kinds of 'showing'/'saying' continua are—from a synchronic and diachronic point of view—inherently problematic. At what point, for example, does a natural sound become a word (or a spontaneous gesture become integrated into a sign language)? How can there be any *continuum* between the analog and the digital? I think the continuum presented in Section 3 between those cases in which the evidence provided for the first layer of information is direct—natural—to those in which it is all indirect may be instructive here. Although this 'showing'/'saying' continuum is orthogonal to the continua proposed by Kendon and Goffman[47] it may shed some new light on them.

In Wharton (2000, forthcoming) I argue that the continuum has evolutionary implications too, and might provide a framework in which we can flesh out an account of the evolution of the metarepresentational abilities required for full-blown inferential communication, and perhaps even the evolution of language.

Because although I've been comparing the meaning conveyed by *natural* signs, *natural* behaviours and biologically-inherited *natural* codes on the one hand, with language on the other, that should not be taken as implying that there is anything *un*natural about the human linguistic code. In fact, the view that there *is* something unnatural about it is somehow implicit in previous approaches to language—a few of which I have mentioned here—that stress its *arbitrary* or *conventional* qualities. I believe this view is mistaken. The very shape of language itself is constrained by a *biologically-inherited*, evolved cognitive mechanism. Acquisition of language is, to humans at least, the most 'natural' thing in the world; it is not learned, it grows (Chomsky 1988: 134). Language is, though admittedly in a different sense, a 'natural' code too.

It's ironic—given that he himself saw language as a system of conventions—that Grice should have provided a possible solution. The central theme of 'Meaning' was that the notions of meaning and intention are inextricably intertwined. Putting this in more current, cognitivist terms, we might say that the 'non-natural' aspects of the human linguistic code are a product of 'natural' human cognitive abilities. In fact,

---

[47] In Wharton (forthcoming) I attempt to collapse Goffman's continuum with the relevance theory 'directness' continuum. I now suspect there are two separate, yet related continua.

Grice himself (1982) provides a sketch of how meaning$_{NN}$ might have arisen from meaning$_{N}$ (1989: 292-297). Thus, Peirce's "[*arbitrary*/*conventional*] symbol-using mind" is re-interpreted as a mind that has the ability to reflect on the content of signals, and the ability to reflect on the intentions behind instances of their use. In fact, it could be argued that the two abilities are fundamentally the same.

At this point we seem to have come full circle, back to the semantic notions that run so strongly through Grice's work. But at least, so construed, non-natural doesn't mean "unnatural", and that, I think, means we're on the right track.

References

Allen, C. & M. Bekoff (1997) *Species of Mind*. Cambridge, MA: MIT Press.

Austin, J. (1962) *How to Do Things with Words*. Oxford: Clarendon Press.

Bach, K. & R. Harnish. (1979) *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., A. Spitz & P. Cross (1993) Do children with autism recognise surprise? A research note. *Cognition & Emotion* 7(6): 507-516.

Blakemore, D. (1987) *Semantic Constraints on Relevance*. Oxford: Blackwell.

Bolinger, D. (1983) The inherent iconism of intonation. In Haiman, J. (ed.). *Iconicity in Syntax* (Typological studies in language, Volume 6). 97-109. Amsterdam/Philadelphia: John Benjamins.

Bradbury, J. & S. Vehrencamp. (1998) *Principles of Animal Communication*. Sunderland, MA: Sinauer Associates.

Carston, R. (1998) *Pragmatics and the Explicit/Implicit Distinction*. University College London PhD Thesis. (To be published by Blackwell: Oxford.)

Chomsky, N. (1986) *Knowledge of Language*. London: Praeger.

Chomsky, N. (1988) *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

Chomsky, N. (1995) *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, N. (2000) *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Clark, H. (1994) *Using Language*. Cambridge: Cambridge University Press.

Darwin, C. (1998) *The Expression of the Emotions in Man and Animals*. London: Harper Collins. Originally published (1872). London: John Murray.

Davies, M. (1996). Philosophy of language. In Bunnin, N. & E.P Tsui-James (eds.). *The Blackwell Companion to Philosophy*. 90-139. Oxford: Blackwell.

Davis, S. (1991) *Pragmatics: A Reader*. Oxford: Oxford University Press.

Duchenne, G. (1990) *The Mechanism of Human Facial Expression*. (Edited and translated by R. A. Cuthbertson). Cambridge: Cambridge University Press. Originally published (1862) as *Mecanisme de la Physionomie Humaine*.

Eibl-Eibesfeldt, I. (1972) Similarities and differences between cultures in expressive movements. In Hinde, R. (ed.). *Non-Verbal Communication.* Cambridge: Cambridge University Press.

Ekman, P. (1989) The argument and evidence about universals in facial expressions of emotion. In Wagner, H. & A. Manstead (eds.). *Handbook of Social Psychophysiology*. 143-164. New York: Wiley.

Ekman, P. (1992) An argument for basic emotion. *Cognition and Emotion* 6: 169-200.

Ekman, P. (1994) Strong Evidence for Universals in Facial Expressions: A Reply to Russell's Mistaken Critique. *Psychological Bulletin*, Volume 115, No. 2: 268-287.

Ekman, P. (1999) Emotional and conversational nonverbal signals. In Messing, L. & R. Campbell (eds.). *Gesture, Speech and Sign.* Oxford: Oxford University Press.

Feibelman, J. (1960) *An Introduction to Peirce's Philosophy.* London: Allen & Unwin.

Fodor, J. (1975) *The Language of Thought.* New York: Crowell.

Fodor, J. (1981) *Representations*. Hassocks: Harvester Press.

Fodor, J., Fodor, J. & M. Garrett (1975) The psychological unreality of semantic representations. *Linguistic Enquiry*, Fall 1975: 515-531.

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.

Gazzaniga, M. & C. Smiley. (1991) Hemispheric mechanisms controlling voluntary and spontaneous facial expressions. *Journal of Cognitive Neuroscience* 2: 239-245.

Goffman, E. (1981) *Forms of Talk*. Oxford: Blackwell.

Grice, H. P. (1957) Meaning. *Philosophical Review* 66: 377-88. Reprinted in Grice (1989). 212-223.

Grice, H.P. (1969) Utterer's meaning and intentions. *Philosophical Review* 78: 147-177. Reprinted in Grice (1989). 86-116.

Grice, H. P. (1975) Logic and conversation. William James lectures. Reprinted in Grice (1989). 22-40.

Grice, H. P. (1982) Meaning revisited. In Smith, N. V. (ed.). *Mutual Knowledge*. London: Academic Press. Reprinted in Grice (1989). 283-303.

Grice, H. P. (1989) *Studies In The Way of Words.* Cambridge MA: Harvard University Press.

Happe, F. (1993) Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition* 48: 101-19.

Happe, F. (1994) *Autism: An Introduction to Psychological Theory*. Cambridge MA: Harvard University Press.

Hauser, M. (1996) *The Evolution of Communication*. Cambridge, MA: MIT Press.

Hockett, C. (1959) Animal 'languages' and human language. In Spuhler, J. N. (ed.). *The Evolution of Man's Capacity for Culture*. 32-39. Detroit: Wayne State University Press.

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.

Ifantidou-Trouki, E. (1993) Sentential adverbs and relevance. *Lingua* 90.1: 65-90.

Jakobson, R. & M. Halle (1956) *Fundamentals of Language.* The Hague: Mouton.

Kaplan, D. (1989) Afterthoughts. In Almog, J., J. Perry & H. Wettstein (eds.). *Themes from Kaplan.* 481-563. New York: Oxford University Press.

Kendon, A. (1988) How gestures can become like words? In Poyotas, F. (ed.). *Cross-Cultural Perspectives in Nonverbal Communication*. 131-141. Toronto: Hogrefe.

Kraut, R. & R. Johnston (1979) Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology* 42: 1529-1553.

Leslie, A. (1987) Pretense and representation: The origins of "theory of mind". *Psychological Review* 94: 412-426.

Leslie, A. & F. Happe (1989) Autism and ostensive communication: The relevance of metarepresentation. *Development and Psychopathology* 1: 205-12.

Levinson, S. (1989) *Pragmatics*. Cambridge: Cambridge University Press.

McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

Millikan, R. (1984) *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.

Okanoya, K. (forthcoming) Sexual Selection Can Shape Finite-State Syntax in Birdsong: A Sexual Display Origin of Syntax. Ms. Chiba University.

Peirce, C. (1987, 1903) Logic as semiotic: the theory of signs. In Buchler, J. (ed.). *The Philosophical Writings of Peirce*. (1955) New York: Dover Press.

Quirk, R., Greenbaum, S., Leech, J. & J. Svartik (eds.) (1985) *A Comprehensive Grammar of the English Language*. London: Longman.

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.

Russell, J. (1994) Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. *Psychological Bulletin*, Volume 115: 102-141.

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. & A. Leslie (1999) Modularity, development and "theory of mind". *Mind and Language* 14: 131-153.

Seyfarth, R., Cheney, D. & P. Marler (1980) Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Animal Behaviour* 28: 1070-1094.

Seyfarth, R. & D. Cheney (1992) Meaning and mind in monkeys. *Scientific American*, Vol. 267/6: 78-86.

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.

Shannon, C. & W. Weaver (1949) *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.

Smith, N. & I. Tsimpli (1995) *The Mind of a Savant.* Oxford: Blackwell.

Sperber, D. (1994) The modularity of thought and the epidemiology of representations. In Hirschfeld, L. & S. Gelman (eds.). *Mapping the Mind: Domain Specificity in Cognition and Culture*. 39-67. Cambridge: Cambridge University Press.

Sperber, D. (2000) Metarepresentations in an evolutionary perspective. In Sperber, D. (ed.). *Metarepresentations: A Multidisciplinary Perspective*. (Vancouver Studies in Cognitive Science 10). 117-139. New York: Oxford University Press.

Sperber, D. (2001) An evolutionary perspective on testimony ad argumentation. To appear in a special issue of *Philosophical Topics* in honour of Alvin Goldman.

Sperber, D. & D. Wilson. (1986/1995) *Relevance: Communication and Cognition.* Oxford: Blackwell.

Sperber, D & D. Wilson. (1998) The mapping between the mental and public lexicon. In Carruther, P. & J. Bouchers (eds.). *Language and Thought: Interdisciplinary Themes*. 184-200. Cambridge: Cambridge University Press.

Strawson, P. (1964) Intention and convention in speech acts. *Philosophical Review* 73: 439-460.

Wharton (2000) Interjections, evolution and the 'showing'/'saying' continuum. Paper presented to The Evolution of Language Conference. Paris, April 2000.

Wharton, T. (submitted) Interjections, language and the showing/saying continuum.

Wierzbicka, A. (1994) Cognitive domains and the structure of the lexicon: The case of emotions. In Hirschfeld, L. & S. Gelman (eds.). *Mapping the Mind: Domain Specificity in Cognition and Culture*. 431-453. Cambridge: Cambridge University Press.

Wierzbicka, A. (1996) *Semantics: Primes and Universals*. Oxford: Oxford University Press.

Wierzbicka, A. (2000) The Semantics of interjection. *Journal of Pragmatics* 18: 159-192.

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 Sperber, D. (ed.). *Metarepresentations: A Multidisciplinary Perspective*. (Vancouver Studies in Cognitive Science 10). 117-139. New York: Oxford University Press.

Wilson, D. & D. Sperber. (1988) Mood and the analysis of non-declarative sentences. In Dancy, J., J. Moravcsik & C. Taylor (eds.). *Human Agency: Language, Duty and Value*. Stanford, CA: Stanford University Press.

Wilson, D. & D. Sperber. (1993) Linguistic form and relevance. *Lingua* 90.1: 1-25.

Wilson, D. & D. Sperber. (2000) Truthfulness and Relevance. *UCL Working Papers in Linguistics* 12: 215-255.