

Agreements, Arguments and Misunderstandings

Benjamin Finn

‘In America I shoot robins and find them thrushes’

George H. Kingsley: *Notes on Sport and Travel* (1900), p.182

Quine, Schiffer and others have tried to demolish the possibility of a theory of meaning; this dissertation attempts to show that the existence of such a theory wouldn't help the practical use of language much in any case. Spurious agreements and misunderstandings, which are caused by various kinds of ambiguity and can occur anywhere, are difficult (and sometimes impossible) to distinguish from genuine agreements and disagreements. This causes problems for translation, communication and philosophical debate.

EXCHANGES

A short story:

Amy, who is American, and her friend Barry, who is British, are out walking in the countryside. Spying a bird, Amy exclaims: “That’s a robin!”

“No, it isn’t,” replies Barry.

Who’s right? At first glance it seems that in this exchange one (and only one) of them must be right; however, little do they know that whereas in Britain ‘robin’ means any bird of genus *Erithacus*, to Americans ‘robin’ denotes the red-breasted thrush, *Turdus migratorius*. Firstly, this means that what seemed like a genuine disagreement was in fact just a misunderstanding, for Amy and Barry speak different dialects. Although they were asserting and denying the sentence ‘That’s a robin’, Amy was asserting the proposition (That’s a robin_A)¹ and Barry was denying the proposition (That’s a robin_B).

Secondly, instead of one (and only one) of Amy and Barry being right, we now have three possibilities: either the bird was actually a robin_A (so *both* were right, in their dialects), it was a robin_B (so *neither* was right), or it was neither (so *only Barry* was right). Rather surprisingly, it is impossible for only Amy to be right – so there is an asymmetry between Amy and Barry. This is because we have omitted a fourth permutation which cannot happen: the case in which the bird is a robin_A *and* a robin_B. (If this permutation could and did occur, Amy would be right and Barry wrong.) But the three legitimate permutations are symmetrical with respect to the dialects, so the asymmetry arises merely from the fact that Amy is making an assertion and Barry is making a denial. (When a single dialect is in use this distinction does not arise, for asserting/denying that *p* is equivalent to denying/asserting that $\neg p$.)

Let’s clarify some terms we have used:

Interpretation: a proposition assigned to a sentence.

Dialect: an assignment of interpretations to sentences.

Agreement: a dialogue in which both participants effectively agree that a sentence expresses a true proposition.

Disagreement: a dialogue in which one participant (the *asserter*) effectively claims that a sentence expresses a true proposition, and the other (the *denier*) effectively claims that it does not.

¹(...) encloses propositions. Subscripts A and B distinguish the American and British senses of ambiguous words. ‘...’ encloses sentences, which express propositions. encloses statements, which are utterances of sentences. The contents of “...” and ‘...’ need not be in English.

Exchange: an agreement or disagreement.

Exchanges needn't be fully explicit, so we can regard the following situation as an 'effective' agreement:

Amy: "That's a robin!"

(Barry thinks she is right, but doesn't reply.)

Similarly, reading text which you agree or disagree with counts as an exchange with the author. The very general nature of exchanges makes them one of the main building-blocks of communication, though there are other important ones (such as questions).

Lying always affects exchanges in an obvious and unexciting way – by negating truth-values of propositions – so we can assume for convenience that any statements made are sincere. Similarly, we need only consider two-person exchanges because ones involving more participants are equivalent to exchanges between pairs of them. For instance:

A: "The earth's flat."

B: "No, it's round."

C: "No, it's an oblate spheroid."

is equivalent to A *versus* B, A v. C and B v. C.

Ten types of exchange

We often assume that if we can interpret another person as speaking English, then their words mean what we think they do – our interpretation of their sentence is the same as theirs. When this is the case, there are two possible types of agreement: one in which both participants are right, and one in which neither is. If the participants disagree, however, then only one of them can be right. Writing 'A' or 'D' for 'agreement' or 'disagreement', '=' for 'same interpretation' and '0', '1' or '2' for neither, one or both participants right, we can abbreviate these cases by the triples $\langle A, =, 2 \rangle$, $\langle A, =, 0 \rangle$ and $\langle D, =, 1 \rangle$ respectively.

However, as in the robin example, the participants may speak different dialects, and so what looks like a genuine exchange could in fact be spurious and rather pointless. This multiplies the number of possible exchanges considerably, especially as the participants may not only be misunderstanding each other, but may be wrong as well! First, some more terminology:

Genuine exchange: one in which the participants' interpretations are the same.

Spurious exchange: one in which the participants' interpretations differ (as in the robin example).

Argument: a genuine disagreement.

Misunderstanding: a spurious disagreement. (Even though spurious agreements also involve a failure of understanding, we should reserve 'misunderstanding' solely for disagreements, in line with everyday usage. English does not possess a single word for a spurious agreement, probably because people seldom realise that such things occur.)

See Table 1, which lists all possible types of exchange. $\langle D, \neq, 1 \rangle$ has been split into $\langle D, \neq, a \rangle$ and $\langle D, \neq, d \rangle$ because in some exchanges (such as the robin example) the denier can be right but the asserter can't. Notice just how devious some of the other types are, particularly $\langle A, \neq, 2 \rangle$ in which the participants agree with each other and are both right, but they are talking about different things. For example:

Amy says "That's a robin", referring to a robin_A sitting on a branch.

Barry replies "Yes, it is," referring to a robin_B sitting next to it.

Agreement/ disagreement	Interpretations same/different	Who's right	Abbreviation	Description
Agreement	Same	Both	<A,=,2>	Genuine agreement
Agreement	Same	Neither	<A,=,0>	Genuine (but misleading) agreement
Agreement	Different (≠)	Both	<A,≠,2>	Spurious agreement
Agreement	Different	One	<A,≠,1>	Spurious agreement
Agreement	Different	Neither	<A,≠,0>	Spurious agreement
Disagreement	Same	One	<D,=,1>	Argument
Disagreement	Different	Both	<D,≠,2>	Misunderstanding
Disagreement	Different	Asserter (a)	<D,≠,a>	Misunderstanding
Disagreement	Different	Denier (d)	<D,≠,d>	Misunderstanding
Disagreement	Different	Neither	<D,≠,0>	Misunderstanding

Table 1: Types of exchange

Disagreements on matters of opinion are characterized by <D,≠,2>; for instance, if ‘That was a delicious meal’ expresses for each of two participants the proposition-schema (I found that meal delicious), they could disagree while both being right. (I.e. the sentence is ambiguous, though its statement isn’t.) Because the ‘I’ is not explicit, it appears as if the participants are expressing the same proposition, so disagreements like this tend to be irresolvable. *De gustibus non est disputandum*.

Three permutations have been omitted from the table – <A,=,1>, <D,=,0> and <D,=,2> – which are impossible for obvious reasons.

ARCHIPELAGO MODEL OF LANGUAGE

Kripke (1982, ch. 2) makes it clear that the possibility of misapplication prevents a speaker’s actual applications of a word from determining its meaning. Moreover, misapplications of a word resemble correct applications of the same word with a different meaning. For instance, ‘68 + 57 = 5’ could be an instance either of incorrect addition or of correct ‘quaddition’. *Being wrong resembles being foreign*, if by ‘foreign’ we mean speaking a significantly different dialect.

Truth and meaning are like two variables in an equation: you can’t solve for both simultaneously without more information. Unfortunately, we do not usually realize this; when someone makes an incongruous statement, we either assume they’re speaking the obvious dialect but are mistaken, or else that they are right but foreign. If this were the case then determining what someone meant and whether they were right wouldn’t be so difficult. However, there is always the possibility that *we* are wrong, or that our opponent is foreign *and* wrong, which can sometimes resemble speaking standard English and being right! For instance, ‘68 + 57 = 125’ might be erroneous *quaddition*.

Now, I use the verb ‘resemble’ with caution above; I’m not suggesting that being wrong *is the same as* being foreign. We shall assume that there is *some* theory of meaning, and this would of course determine the fact of whether a sentence means what we think it does. However, we might not know this fact – as in the robin example, an innocent-looking sentence might *seem* like our dialect but have an unexpected meaning.

This seemingly contrived example matters because of the central argument of this dissertation: we shall argue that each English-speaker possesses a personal dialect, or *idiolect*. These overlap considerably to produce a language which permits communication, but leaves plenty of scope for misunderstandings, especially hard-to-detect or even undetectable ones. This entails that meaning is something in between being communal and being private – it is public, but individual.

A bit more story:

Charles (who is also British) is walking alongside Amy and Barry. Inspired to poetic language, he exclaims: “The grass looks vermeil!” Barry says, “No, you’re wrong – it looks *verdant*.”

‘Vermeil’ means ‘bright red’ in English, so it looks like this is an ordinary argument (<D,=,1>) in which Charles is wrong. Let’s suppose that:

(i) Charles believes ‘vermeil’ means ‘green’ in English.

We can make the following weak assumptions:

(ii) Charles knows what he’s *attempting* to say,

(iii) Charles means his statement sincerely,

(iv) Charles is not hallucinating or insane.

Let’s call Charles’ statement *s*, and say that he is trying to express the proposition that *p*. (iv) rules out the possibility that Charles doesn’t believe that the grass looks green, so it follows uncontentionally from (i) and (iii) that $p = (\text{The grass looks green})$. Let’s propose, contentiously, that

(v) *s* means (The grass looks green).

Objector: *But under community views of meaning, Charles’ statement means (The grass looks bright red), just in virtue of the fact that ‘vermeil’ means ‘bright red’ in English and Charles speaks English. So s does not express p at all – it is a misarticulation of p.*

However, I propose that the following requirements are sufficient for *X*’s statement *s* to express *p* properly:

(a) *X* utters *s*.

(b) *X*, who speaks dialect *L*, believes that this utterance of *s* expresses *p* in his dialect. (This rules out sleep-talking, in which it is not clear that statements like “I want to go to Elbow” express anything much.)

(c) Understanders of *L* who hear *s* take it to be expressing *p*.

What more could one desire of statements than that they convey propositions between speakers of a dialect *salva signification*? These requirements rule out genuine misarticulations; for instance, if Charles knew that ‘vermeil’ means ‘bright red’, but only said it (rather than ‘verdant’) due to a slip of the tongue, (b) and (c) would fail.

However, (i) and (ii) imply that Charles believes that *s* expresses *p*, thus fulfilling (b). Now to fulfil (c): the one person we can be sure speaks Charles’ dialect *L* is none other than Charles himself. Given (i), if we were to tape-record *s* and play it back to Charles at a later date, he would understand it as expressing *p*.

(a), (b) and (c) are fulfilled, so *s* *did* express *p* after all. From this (v) follows immediately, so Charles’ statement was, in fact, *true*. However, interpreted in standard British English (as found in dictionaries), *s* meant (The grass looks bright red), which is false. Hence, Charles was speaking a dialect which differs from standard English at least in the meaning of ‘vermeil’. Consequently, the exchange was of type <D,≠,2> – a misunderstanding. Perhaps, then, a more correct (and more diplomatic) reply for Barry to have given would have been:

“Yes, you’re right – but the English word for that colour is ‘verdant’.”

Three types of mistake

Just what sort of mistake is Charles making, since it isn’t a fallacy (about the colour of the grass) or a misarticulation? He is mistaken about the meaning of the word ‘vermeil’ in the dialect he is *trying* to speak, namely standard British English. However, since he doesn’t explicitly say “I believe that in

standard English ‘vermeil’ means ‘green’”, his statement’s truth-value isn’t affected by the mistake.² We can call Charles’ error – being mistaken about the correct meaning of a word – a *malapropism*. So we now have three types of mistake: a *fallacy*, which renders statements false; a *misarticulation*, which renders them incomprehensible; and a *malapropism*, which tends to cause spurious exchanges.

Very few of us know accurately – or at least with accuracy sufficient to get the truth-conditions right – the dictionary definition of every word we use. Any exchange in which we commit a malapropism shows that we, like Charles, don’t speak standard English. Moreover, as it is unlikely that any two of us make precisely the same malapropisms, our *dialects* all differ. (Even if two people did have identical dialects, the fact that they couldn’t know this when talking to each other makes them just as linguistically isolated as anyone else. Thus even they effectively speak idiolects.)

Conclusion 1: *We all speak idiolects.*

I am not suggesting that our idiolects are independent, otherwise communication would be impossible; we do all ‘speak English’. ‘Speaking English’ can’t, however, mean speaking standard British (or American) English, for virtually none of us do that. It must instead mean speaking a dialect which is sufficiently close to the standard form to make communication with other ‘English speakers’ possible.

Wittgenstein’s Rule-Following Argument shows that meaning *p* by *s* isn’t constituted by following the correct pattern of application;³ it seems that speaking a certain language isn’t constituted by using the correct set of meaning-assignments, either. The language only specifies ideal meanings for sentences; speakers can and do deviate from these a certain amount while still ‘speaking the language’. Hence, counterintuitively,

Conclusion 2: *The language someone speaks and the meaning of a statement of theirs in that language don’t determine the meaning of that statement.*

For instance, Charles speaks English, and ‘The grass looks vermeil’ means (The grass looks bright red) in English, but Charles’ statement didn’t mean (The grass looks bright red).

The community, then, is like an archipelago (a cluster of islands): the speaker’s dialects are close but distinct. Although we haven’t presupposed any particular theory of meaning, this forces us to abandon at least some community-based accounts; however, it doesn’t rule out externalism in general.

Two corollaries:

Colour-blindness

Red-green colour-blind people in Charles’ situation similarly cannot be mistaken about how the grass looks *to them*, but the general pattern of what things they call ‘vermeil’ as opposed to ‘verdant’ will be abnormal.⁴ Some green shades they will frequently call ‘vermeil’, and some red shades they will frequently call ‘verdant’. Consequently, their idiolects will be much more similar to each other’s than to normal people’s. It may not be too extreme, therefore, to claim:

Conclusion 3: *Colour-blind people speak a different language from everyone else.*

Language development

A comparison of the definitions of ‘infer’ in British and American dictionaries reveals a difference in meaning: ‘infer_A’ can not only be used in the British sense ‘deduce’, but also to mean ‘imply’. Let’s suppose ‘infer’ originally only bore the British sense, but this changed to ‘infer_A’ in America later. Under a community view this would be problematic, for the use meaning ‘imply’ could have been correct only when enough people were using it in America for it to be a *bona fide* new sense. Statements like “Whenever people use sarcasm, they infer the opposite of what they say” would have

²This contrast between explicit and implicit falsehoods resembles that between “I will pay you sixpence tomorrow” and “I promise I will pay you sixpence tomorrow” when unfulfilled, discussed in Austin (1975).

³See Kripke (1982), ch.2.

⁴Cupitt (1979, p. 93) puts forward the idea of colour-blindness as a linguistic handicap.

Type of ambiguity	Constraint on truth-conditions	Types of exchange caused by the ambiguity			
Contradictory interpretations	$\neg(p \equiv q)$	$\langle D, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$		$\langle D, \neq, 2 \rangle$
q includes p	$p \rightarrow q$	$\langle A, \neq, 0 \rangle$ $\langle D, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$ $\langle D, \neq, a \rangle$	$\langle D, \neq, d \rangle$	$\langle A, \neq, 2 \rangle$
p includes q	$q \rightarrow p$	$\langle A, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$ $\langle D, \neq, a \rangle$	$\langle D, \neq, d \rangle$	$\langle A, \neq, 2 \rangle$ $\langle D, \neq, 2 \rangle$
Incompatible interpretations	$\neg(p \& q)$	$\langle A, \neq, 0 \rangle$ $\langle D, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$	$\langle D, \neq, d \rangle$	$\langle D, \neq, 2 \rangle$
Jointly exhaustive interpretations	$p \vee q$	$\langle D, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$ $\langle D, \neq, a \rangle$		$\langle A, \neq, 2 \rangle$ $\langle D, \neq, 2 \rangle$
Independent interpretations	none	$\langle A, \neq, 0 \rangle$ $\langle D, \neq, 0 \rangle$	$\langle A, \neq, 1 \rangle$ $\langle D, \neq, a \rangle$	$\langle D, \neq, d \rangle$	$\langle A, \neq, 2 \rangle$ $\langle D, \neq, 2 \rangle$

Table 2: Types of ambiguity

been just plain false as uttered by an British speaker – sarcasm isn’t a form of inference! However, when so many people were making statements like this that this use of ‘infer’ was valid, then the same statement in the same situation would have been true.

Why does a *smooth* change in the number of users of the sentence produce a *sudden* change in truth-value? Perhaps the truth-value became indeterminate during the change-over period from the British to the American dialect. But, in any case, a community theory of meaning would have to explain why the invalid sense of ‘infer’ could be used successfully by someone to communicate at least with himself (in a diary, say).

Under the archipelago model, however, these problems evaporate. The sentence was true, and the ‘infer_A’ sense was valid, *all along*; the infer_A-pioneer’s mistake was the merely malapropistic one of thinking that other English-speakers would understand it.

AMBIGUITIES

Six types of ambiguity

An *ambiguous* sentence is one which different people interpret differently. It can cause spurious exchanges, but not necessarily all the types in Table 1. For instance, we have already seen that the asserter (Amy) cannot be right on her own in a disagreement over ‘That’s a robin’, so this sentence can’t cause $\langle D, \neq, a \rangle$.

However, not all ambiguous sentences behave the same way. If we call the two interpretations of ‘That’s a robin’ p and q , then of the four circumstances $(p \& q)$, $(p \& \neg q)$, $(\neg p \& q)$ and $(\neg p \& \neg q)$, all but the first are possible; that is, the constraint on the truth-conditions of ‘That’s a robin’ is $\neg(p \& q)$. For other sentences, each of the four circumstances could be possible or impossible, giving $2^4 = 16$ different ways of being ambiguous. We can exclude, however, the nine cases in which p or q is either necessary or impossible – as we may assume the exchange concerns something contingent – and also the case in which $p \equiv q$, where there is no ambiguity at all. This leaves us with six types of ambiguity, listed in Table 2. In the case of disagreements, p is what the asserter asserts and q is what the denier denies.

This raises the issue of the Gricean distinction between what’s *literally* said and what is *conveyed* by the saying of it. Austere theories of meaning would rely on supplementation from conversational and situational context; liberal ones would build all of this into the truth-conditions. We are taking propositions as the unit of meaning purely for the convenience it confers to the analysis of exchanges; this decision should be independent of any theory of meaning. For instance, we can allow that Amy’s intonation may lay an otherwise unambiguous statement open to a sarcastic interpretation without

specifying precisely under what circumstances this occurs.

Contradictory interpretations

This type is found most often in misunderstood sarcasm. For instance:

Amy (sarcastically): “Frinton-on-Sea is really lively.”

Barry (not realising Amy was being sarcastic): “No, it isn’t!”

Both participants are unknowingly agreeing with each other. The sarcastic (*p*) and literal (*q*) interpretations of Amy’s sentence cannot be right simultaneously, so Amy and Barry are either both right or both wrong.⁵

Sarcasm is the lowest form of wit; more specifically, it is a weak form of pun between the literal and sarcastic interpretations. Whereas ambiguities are often accidental and undesirable, sarcasm relies on them for its effect.

Inclusive interpretations

We distinguish two types – *q* includes *p*, and *p* includes *q* – because they cause different sorts of exchange in disagreements. When the participants agree, however, they are equivalent – the letters *p* and *q* in the table are just interchanged.

These types of ambiguity usually have one meaning more precise than the other. For instance,

Amy: “The earth’s round.”

Barry: “No, it isn’t” (for it is an oblate spheroid).

Of course, they are both right, since Amy meant *roughly* spherical and Barry pedantically⁶ meant *precisely* spherical; *q* is more precise than *p*.

Incompatible interpretations

This is the ‘robin’ type: it often occurs when the interpretations concern two varieties of the same kind of object. There is an asymmetry betweenasserter and denier – only the denier can be right on his/her own, so the denier is always ‘at least as right’ as the asserter.

An example with an ambiguous verb:

At 5 pm precisely, Barry tripped along the path to his house.

‘Tripped’ means ‘pranced’ or ‘stumbled’; Barry could have either pranced or stumbled, but not both simultaneously. Note that ‘At 5 pm precisely’ is necessary – without it, the sentence would be true in *both* senses had Barry pranced at 5 pm and stumbled at 5.01 pm, so that would be a case of independent interpretations.

Jointly exhaustive interpretations

Here, it is always the case that one interpretation or the other is true; e.g. ‘That’s *not* a robin’ is bound to be true under either the British or the American interpretation. As with the previous ambiguity type there is an asserter/denier asymmetry, though this time it is biased in the asserter’s favour. However, the types are effectively equivalent when it comes to disagreements, since:

Barry: “That’s not a robin.”

Amy: “You’re wrong – it *is* a robin.”

is equivalent to:

⁵Ignoring the fact that by being sarcastic Amy is asserting that (Frinton-on-Sea is really unlively) rather than the banal ¬(Frinton-on-Sea is really lively). This means that there is a case in which Amy and Barry are both wrong, *viz.* if Frinton-on-Sea is neither particularly lively nor particularly unlively.

⁶I don’t accuse Barry of being pedantic just because he is being precise here, for pedantry often involves imprecision – increasing (rather than restricting) the extension of a word. For example, claiming that a square is a type of rectangle, or that a bee is an animal, is pedantic by using ‘rectangle’ and ‘animal’ *less* precisely than usual.

Amy: “That’s a robin.”

Barry: “No, it isn’t.”

(with the statements reworded and swapped).

Independent interpretations

This case often occurs when the interpretations deal with two unconnected topics. *Double entendres* are usually of this type; as with sarcasm, their comic effect relies on there being two possible interpretations. (Perhaps there are further kinds of pun corresponding to the other types of ambiguity.)

Meaning and ambiguity depend on context

In the above description of ambiguities, we have been glossing meaning as truth-conditions; is this appropriate, or should we consider ambiguities between even finer shades of meaning? Although we can recognise a *poetic* difference between (say) ‘that grass is green’ and ‘that grass is verdant’, it isn’t enough to create a possible difference in truth-value. But possible difference in truth-value *just is* difference in truth-conditions, so the coarse grain of meaning provided by truth-conditions is exactly what we need; subtler distinctions don’t matter.

It seemed with the ‘robin’ case that a difference in meaning between the words used by the participants was sufficient to render an exchange spurious. However, in the following:

Amy: “We can infer from the clouds that it’s going to rain.”

Barry: “That’s right.”

and in many everyday cases between idiolects, there seems to be no possibility of confusion. How can we account for this?

Clearly some conversations between Amy and Barry can make apparent the difference in meaning between infer_A and infer_B ; for instance:

Barry gives Amy *Cooking for Beginners* for Christmas.

Amy: “Are you inferring_[A] that I can’t cook?”

Barry: “Yes – I infer_[B] it from the way you make toast in a saucepan.”

The context here admits two interpretations of ‘infer’, which it didn’t in the previous example. We can deduce that it was wrong to regard ambiguity as a property of individual words like ‘robin’; other aspects of the exchange can restrict the meaning. (One obvious way to do this is to append an adjective, e.g. ‘British robin’.) If an exchange involves two overlapping meanings of a word – such as infer_A and infer_B – the exchange will nonetheless be genuine if the context rules out those aspects of the meanings which are not in common. So although ‘infer’ *on its own* means different things in American and British English, this doesn’t entail that sentences containing ‘infer’ do.

Conclusion 4: *An exchange involving a word which means different things to the two participants is not necessarily spurious.*

Because meaning depends on context, so does synonymy (i.e. genuineness). This has corollaries for two different branches of philosophy:

Incommensurability of theories

Consider the following exchange:

Newton: “This bag of sugar has a rest mass of 1kg.”

Einstein: “That’s right.”

But Newton meant Newtonian mass and Einstein meant Einsteinian mass, so isn’t this a spurious agreement? No – because the truth-conditions of Newton’s sentence are the same under both interpretations. Similarly, there needn’t be a divergence of meaning between the *sentences* of

conflicting theories even if the *words*’ meanings differ.

Philosophical analysis

Let us suppose that in America, unlike in Britain, ‘knowledge’ literally means ‘true justified belief’. Before Gettier examples were invented the difference in meaning between American and British English probably wouldn’t have come to light, and (in line with Conclusion 4) everyday exchanges between Americans and Britons on the subject of ‘knowledge’ would have been genuine. Suppose Amy and Barry start discussing ‘knowledge’ and come across a Gettier example which causes a disagreement over whether or not this is really an instance of ‘knowledge’. This will be a misunderstanding, for Amy and Barry are *now* quite clearly using ‘knowledge’ in two different senses. This was not the case, however, in their previous use of the word, so we are forced to conclude that at least one of Amy and Barry must have shifted meaning a bit. In general,

Conclusion 5: *When we make philosophical analyses of words, they don’t mean the same as in everyday use.*

This is not too surprising, since analysis involves sharpening up meanings in a way which can plausibly affect truth-conditions slightly. However, Conclusion 4 means this does not prevent the philosophical and everyday senses being totally identical *in almost all contexts*.

ESTABLISHING THE EXCHANGE-TYPE

Idioms, malapropisms and fallacies

Suppose some future civilisation discovers the few surviving pages of twentieth-century written English, deciphers them and reads references to ‘the four corners of the earth’. They might be forgiven for thinking that English-speakers believed the earth was flat. Of course, this is just an idiom which resembles a fallacy; more specifically, it resembles a *common* fallacy – one possessed by almost all speakers of the language – since almost all speakers might happily use the phrase.

What’s the difference between an idiom and a fallacy? A fallacy is a deviation from the pattern of correct applications, whereas an idiom is a deviation of a correct application from a *natural* pattern.⁷ (‘Red hair’ is a less natural application of ‘red’ than ‘green grass’ is of ‘green’. We needn’t dwell on precisely what constitutes naturalness here.) Figure 1 illustrates the distinction. If we don’t know the meaning of a word or sentence well enough to know whether its *correct* applications match A or B, then we only have the *actual* applications (white circles) to go on; and we can’t tell from these whether an abnormal-looking application is a fallacy (A) or a correctly-applied idiom (B).

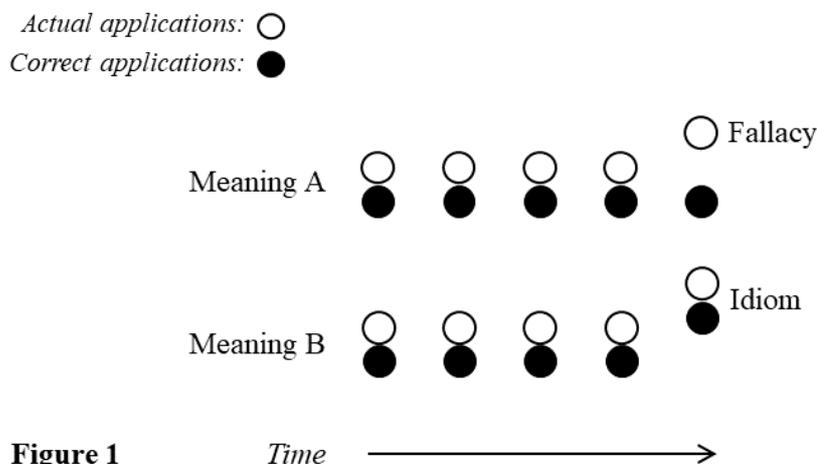


Figure 1

⁷‘Quus’ is an idiom that initially looks like a fallacy (an adding mistake).

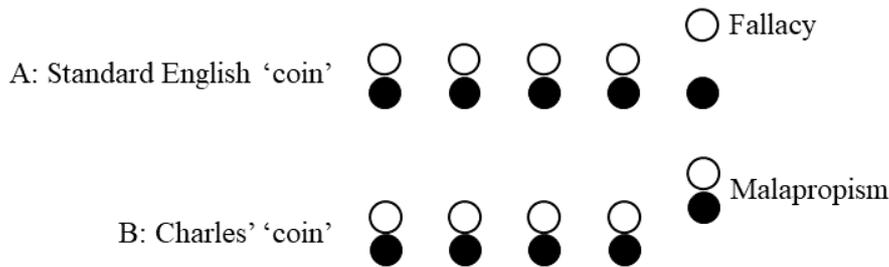


Figure 2

Let's suppose that Charles suffers from the fairly common malapropism of thinking that to 'coin' a phrase means to use one; in fact, it means to invent one. Compare Figure 1 with Figure 2, which illustrates the difference between a malapropism and a fallacy; if we don't know Charles' idiolect well enough to know whether his correct use of a word or sentence follows A or B, we can't tell whether an actual application is a fallacy (A) or a malapropism (B). It is clear from the diagrams that idioms and malapropisms are very alike; the essential difference is that an idiom tends to be common to (virtually) all speakers of a language, whereas a malapropism usually affects just a few people.⁸ Hence idioms resemble *common* fallacies, while malapropisms resemble *individual* ones.

Just as a future civilisation could mistake our idioms for fallacies, some disagreements between English-speakers can be misunderstandings caused by malapropisms. (The 'vermeil' dispute earlier was an example.) Nor are agreements immune: an apparently genuine agreement could be caused by the superimposition of a malapropism and a fallacy (see Figure 3). It could also be caused by A committing a fallacy which matches B's non-fallacious malapropism. Similar cases hold for idioms.

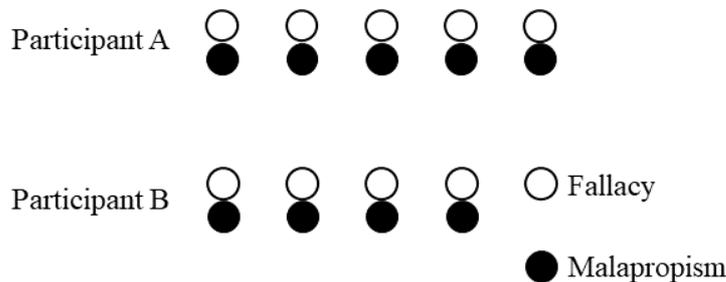


Figure 3

How to detect a spurious exchange?

For communication to be reliable, it is necessary to detect when a message is or isn't getting through; that is, to detect when exchanges are spurious. But in the remainder of this dissertation we will make a case for the following:

It is difficult or impossible in practice to establish whether an exchange with someone else is genuine or spurious. Moreover, it is impossible *in principle* to establish whether a translation of a sentence written in a dead language is correct or not.

An objector will suggest various ways in which a spurious exchange could be detected, which I will respond to.

We shall assume that sentences in languages we are unfamiliar with could contain idioms, malapropisms, or fallacies. For languages we are familiar with (such as English), we can allow, in the objector's favour, that malapropisms are unlikely to create new idioms or destroy existing ones.

⁸This is not always the case. There are common malapropisms, such as the belief that 'enormity' means 'being enormous'. Of course, when malapropisms become common enough they cease to be malapropisms at all.

Excessive charity

Objector: “If we consider a large amount of speech or text in a suspicious dialect, the correct interpretation is the one which renders it the most comprehensible, coherent or truthful. This is roughly the rule-of-thumb we use when guessing the natural language of text; we go for comprehensibility first (to distinguish English from German, say), then maximise coherence or truthfulness (e.g. to choose between the British and American interpretations of ‘Amy put gas in her car’).”

This somewhat resembles Davidson’s ‘Principle of Charity’ – we charitably assume the speech or text must be pretty comprehensible, coherent and truthful, and use this to work out what it meant. Unfortunately, without further restrictions we can be *too* charitable, as is best illustrated by an example:

Charles buys a copy of *The Times* and accidentally starts reading it in French.⁹ He finds that where articles are about anything at all they tend to be concerned a great deal with tea (‘the’), and include many missing accents, spelling errors (e.g. ‘magazine’ misprinted as ‘magasin’, ‘once’ as ‘onze’), and nonsense-words such as ‘building’. Very (*very!*) few sentences are grammatical or coherent. When Charles has realised his mistake, he tries reading the newspaper in other languages, and reveals whole new layers of meaning. Having exhausted the possibilities offered by natural languages, he goes on to devise his own, which make the newspaper much more interesting than in the original English.

Especially when trying to translate an unknown language, what can prevent this sort of gross misinterpretation from happening by accident? So strong is our habit of inferring the language intended that we tend to assume that only one interpretation is possible, but we don’t yet have any proper basis on which to rule out some rather than others. Therefore Charles can devise dialects in which *The Times* reads as the complete works of Franz Kafka, chapter 5 of *Submarine Bambrology* or as ‘ $2 + 1 = 3$ ’. He should thus have little difficulty rendering *The Times* more comprehensible, coherent and truthful than the English version.¹⁰

Conclusion 6: *Maximizing comprehensibility, coherence and veracity doesn’t work.*

Objector: “The newspaper could just state what language it is written in – no inference from the text would be required.”

Unfortunately, it would not then specify in which language the language-name itself was to be understood. (A similar case: suppose an alien were demonstrating his mathematics to humans who did not know which base he was using. He could not clarify this by *stating* the base, since all base-names written in their own base appear as ‘10’.)

Fixing the truth-value

Objector: “Recall the ‘vermeil’ example. We said that whatever sort of mistake Charles was making, it couldn’t be a fallacy. We could tell that it wasn’t a misarticulation by repeating Charles’ statement back to him and confirming that it’s what he meant to say. So his mistake could only have been a malapropism. This determined the disagreement to be of type $\langle D, \neq, 2 \rangle$. Can’t we do something similar in other cases to establish whether an exchange is genuine or spurious?”

Because we could rule out a fallacy in this case, we effectively fixed one variable (the truth-value) in our two-variable equation. This enabled us to solve for the other (the meaning). Uses of sentences in circumstances like this where we can rule out the possibility of a fallacy we shall call *paradigms*. Apparent fallacies uttered in paradigm situations force people to consider the possibility that the

⁹ My brother thought up this idea a few years ago.

¹⁰ Another example is decoding an encrypted message by finding a key which produces plausible text rather than gibberish, such as was done during World War 2 to crack Enigma Machine codes. There is no guarantee that an incorrect decryption won’t be more plausible than the correct one.

speaker is foreign: if someone said “The earth is flat” (not a paradigm case) we would understand him but think *he was wrong*; if, however, he said “That grass looks red”, we would scratch our heads and try to work out *what he meant*. We certainly wouldn’t take the statement at face value.

If we can reduce a situation to a paradigm then we can solve for the meaning. An example: it isn’t clear what shape the medieval traveller Sir John Mandeville thought the earth was.¹¹ He seems both to be aware that the earth is spherical, and to have trouble reconciling this with his other mental picture of places on a flat map. Let’s suppose he claimed “The earth is flat”, and we want to work out what he meant and whether he was right. Establishing that he described objects like table-tops as ‘flat’ is not enough to establish that by ‘flat earth’ he meant ‘earth shaped like a table-top’; perhaps ‘flat’ to him meant *locally* flat – rather as we might describe the surface of the sea as ‘flat’ – or perhaps ‘flat earth’ was to him a poetic idiom, like ‘the four corners of the earth’.

To fix the truth-value of Mandeville’s claim, we need only (!) take him out in a space-ship so he can see the earth, point to it (saying “That is the earth”) and ask if he *now* claims “The earth is flat”. He cannot now plausibly be mistaken about what the earth or its shape are, so we have fixed the truth-value of his sentence as interpreted in *our* language. If he does still claim “The earth is flat”, then he probably means locally-flat or poetically-flat. However, if he now says “I was wrong – the earth is not flat”, then he meant something like our ‘flat’.

We needn’t, of course, find the *actual* truth-value to discover what Mandeville means; all that is required is that both he and us are discussing the same state of affairs. So we could depict a sphere or a plane on a piece of paper, say “Let’s *suppose* the earth looks like this”, and ask him whether he now claims “The earth is flat”. This makes the whole enterprise more practical for cases like the earth’s shape, which are awkward to demonstrate.

Note that some claims Mandeville might make which apparently refute ‘The earth is flat’ may, in fact, not do so; “The earth is spherical”, for instance, wouldn’t refute the original claim had it meant (The earth is locally-flat). *Apparent* refutation isn’t good enough; Mandeville must admit he was wrong for us to fix roughly what he meant. If Mandeville *had* meant literally ‘flat’ but later stubbornly refused to admit he had been incorrect, we would attribute the wrong dialect to him; if we believe Mandeville to be sincere, then at no stage can we demonstrate that his original claim was false. So...

Conclusion 7: *You can avoid being proved wrong by refusing to admit it*

...since there will always be some dialect in which all your claims are true. This little objection aside, reduction to paradigms does seem a sound way of fixing the truth-value, getting a good idea of the meaning and thus establishing whether an exchange is genuine or spurious.

Unfortunately though, two large classes of cases won’t work:

Abstract cases

Paradigms only seem to exist for correspondence-type concepts such as objects and colours (and not necessarily for all of them, either). Statements concerning abstract concepts don’t seem to be reworkable in such a way that they become plainly true. They certainly can’t become *perceptually* unmistakable, otherwise we would be well on the way to performing an empiricist reduction of the whole of language; and *conceptual* unmistakability seems rather a dubious idea. So...

Conclusion 8: *Abstract cases don’t generally reduce to paradigms.*

Names and natural kinds

Under Kripke’s causal theory of names (Kripke 1980) the properties attributed to (say) ‘the Pope’ by the speaker do not determine the referent of ‘the Pope’. If, therefore, Amy said “The Pope lives in Buckingham Palace”, we could not establish by questioning her whether this was a fallacy or whether by ‘The Pope’ she meant the Queen.¹² Neither the Pope nor the Queen nor anyone or anything else

¹¹Mandeville (1983), pp. 192-3.

¹²This is based on a true story: in the summer of 1980, an American tourist got off a coach outside Buckingham Palace and

can count as a paradigm here, since the contents of Amy's mind don't determine what she meant.

Conclusion 9: *Questioning the speaker can't establish whether many exchanges involving names or natural kind terms are genuine or spurious.*

(We should include the 'many' caveat because in some sentences the referent of the name or natural kind term does not affect the truth-value.)

Heavily externalist theories say *no* meanings are determined by the mind of the speaker, with even more unfortunate results. We might, of course, regard Conclusion 9 as a *reductio ad absurdum* for them or for Kripke's causal theory.

Fixing the meaning

Objector: "So using paradigms only sometimes works. How about if instead of fixing the truth-value we fix the meaning, by getting the speaker to reword his/her statement in a dialect which we can be sure both of us understand? If we allow non-verbal communication into languages for this purpose, pointing might be suitable."

We reply with a quote from a poem:

Two voices are there: one is of the deep;
It learns the storm-cloud's thunderous melody,
Now roars, now murmurs with the changing sea,
Now bird-like pipes, now closes soft in sleep:

And one is of an old half-witted sheep
Which bleats articulate monotony,
And indicates that two and one are three,
That grass is green, lakes damp, and mountains steep

– J.K. Stephen: *A Sonnet*, from *Lapsus Calami* (1896)

In this parody of Wordsworth, the 'two voices' of the first and last four lines distinguish two modes of discourse which we shall call high- and low-level.¹³ These are ends of a continuum: lower-level terms are used to define or explain higher-level terms. The lowest level consists of those words which are definitionally most prior – namely the words we are taught first, which denote the simplest concepts. As this initial teaching tends to be done by ostensive definition, it is quite hard to attach the wrong meanings to these words. (Quine (1960, ch. 2) points out that, due to the inscrutability of reference, even ostension could be fairly ambiguous; we can give the objector the benefit of the doubt and assume that we all understand it well enough to avoid significant misunderstandings.) The abstraction required to assemble lower-level concepts to form higher-level ones makes the meaning conveyed by teaching more ambiguous the higher you get, particularly when you reach metaphysical terms.

If someone says something which we suspect we may be misinterpreting, then we can get close to fixing the meaning by getting the speaker to restate his/her sentence at a lower level, thus rendering it less open to misinterpretation. We could get a pretty good idea of what Mandeville's claim meant by asking him to draw a picture of his 'flat earth'. From this we could interpret the truth-value of his utterance "The earth is flat". If he drew Figure 4, then by 'flat' he meant *literally* flat, so he was wrong. If, however, he drew Figure 5, then he probably meant 'locally-flat' or 'poetically-flat', so he was right. (We needn't be able to establish the *precise* meaning in order to work out the truth-value.)

refused to get back on again until he had "seen the Pope".

¹³By analogy with computer languages, in which high-level statements are defined in terms of low-level ones.

Figure 4

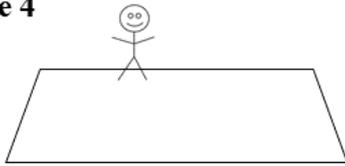
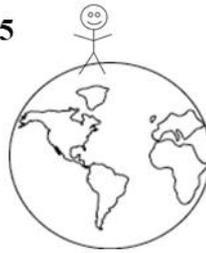


Figure 5



The enormous similarity of this to the paradigm method demonstrates that using paradigms and low-level discourse are really two sides of the same coin – for paradigms occur at the lowest level. We can use the word *reduction* for both, not only because ‘reducing’ the level of discourse from high to low is involved, but also because it corresponds quite well to the philosophical sense.

The problems with reducing to paradigms carry over to low-level discourse: names and natural kinds still won’t work. Abstract concepts, though possibly quite low-level in the way we have defined it, are far from ostension and thus open to enormous misinterpretation.

Objector: *“Is it necessary to reduce at all? Can’t we establish the meaning of a high-level sentence by rewording using other high-level terms?”*

Probably not. Consider how high-level terms such as ‘unethical’ are learned: perhaps ‘should’ was originally taught by abstraction from practical cases, and less common terms such as ‘obligation’, ‘immoral’ and ‘unethical’ were defined in terms of this. Ambiguity would most likely have arisen at the abstraction stage, and the connection between the terms thereafter is fairly analytic. So going round a circle of interrelated abstract terms probably won’t reveal anything about the meaning of the suspicious term at all.

Conclusion 10: *Rewording without reducing is not a reliable way of clarifying meanings.*

Idiom-infested languages

Objector: *“Surely we can identify an idiom without reducing to the sub-verbal level of ostension, using the following verbal trick:*

Françoise (who is French) is walking alongside Barry. Françoise is not aware of the ‘red hair’ idiom in English.

Barry: “Amy has red hair.”

Françoise: “No, she doesn’t” (for Amy’s hair is ginger, not literally red!)

How can Françoise establish whether Barry is uttering an idiom, a fallacy or both? She thinks for a bit, then asks:

“Is Amy’s hair the same sort of colour as blood?”

Barry: “No.”

By circumlocuting the questionable phrase ‘red hair’, Françoise has established that Barry’s original sentence was an idiom; had ‘red’ been unidiomatic, Barry would have replied ‘Yes’.”

But this linguistic trick would have failed if (unknown to *Françoise*) every possible circumlocution of ‘Is Amy’s hair red?’, even those which didn’t mention ‘red’ at all, suffered from a similar idiom so contrived that Barry always correctly replies “Yes”. For instance, if ‘same colour’ suffered from the idiom that it relates ‘red’ hair and genuinely red objects, then Amy’s hair would be ‘the same colour as’ blood. A language which behaves in this inconvenient way we shall call *idiom-infested*.¹⁴ In an idiom-infested language, no amount of questioning can distinguish an idiom from a fallacy;

¹⁴The quus-follower’s idiolect is an example.

reduction to a sub-verbal level (such as pointing at an actual sample of hair) is required.

Objector: *“But it is grossly implausible that any real language would be idiom-infested.”*

English is virtually such a language. We call light and dark blue ‘blue’ even though they look fairly different. It is an idiom of English that we can’t similarly refer to ‘light red’ and ‘dark red’ – meaning *pink* and red – even though the colours stand in the same physical and visual relation to each other as light and dark blue do. The concept of ‘light red’ seems nonsensical; we think of red, try to make it lighter, and give up when we reach any colour we wouldn’t *call* ‘red’.

Now suppose Françoise shows Barry a swatch of light blue and one of dark blue, and asks if they are “the same sort of colour” (or some other circumlocution which doesn’t mention ‘blue’). He may well reply “Yes” – for they are both blue! However, suppose he is shown red and pink, and asked the same question; he might well be forgiven for replying “No” – for, after all, they are not both red! Even though there is no direct mention of the colour-words, perhaps Barry cannot help doing a *de dicto* rather than a *de re* comparison.

In a minor variant of English, Françoise’s earlier attempt to establish that ‘red hair’ was an idiom would fail:

“Is Amy’s hair the same sort of colour as blood?”

Barry: “Yes” – for both are red!

Attempts to uncover the idiom by asking whether or not Amy’s hair was ‘*really* red’, or via any other such circumlocutions, could fail in a similar way. Françoise would be forced to reduce to a level lower than speech to prevent *de dicto* comparisons; for instance, she could point at Amy’s hair and ask Barry if he *now* claimed it was ‘red’. Because Barry could not then reasonably be mistaken about the colour of Amy’s hair, his reply would reveal the ‘red hair’ idiom.

Conclusion 11: *Because languages can plausibly be idiom-infested, idioms in unfamiliar languages cannot reliably be spotted by reduction to any but a sub-verbal level.*

Objector: *“Surely it would be obvious to a speaker when he was using an idiom, particularly if his language was idiom-infested – couldn’t he resolve any disagreement just by pointing this out?”*

The case we have just seen suggests that a single idiom such as ‘red hair’ can spawn a plethora of natural-seeming idioms if speakers of a language make judgements in a *de dicto* rather than a *de re* manner; and this does seem to happen much of the time. So would even the originating idiom seem like an idiom? Again, not necessarily: it doesn’t particularly strike users of the word ‘pink’ how strange it is that we don’t call it ‘light red’.

Conclusion 12: *Not even an idiom-infested language need seem idiomatic to the speaker.*

OUTCOME

At this point the objector runs out of objections, so we can put the various ingredients together to produce our final conclusions:

Translating dead languages

Consider text written in a dead language. Names and natural kinds are untranslatable in principle to start with (Conclusion 9). There is no reason to think the language doesn’t have plenty of idioms lurking which could well crop up inconspicuously at any moment. We can’t, however, ensure correct translation by applying a modified Principle of Charity (Conclusion 6); nor will discovering non-reductive rewordings of suspicious text help (Conclusion 10). Conclusion 11 says a verbal reduction, such as could be provided by context, explanations in the text or contemporary dictionaries, just isn’t good enough for showing up idioms; reduction for abstract terms couldn’t work in any case (Conclusion 8). Because the language is dead, we can’t discover its idioms by performing a sub-

verbal reduction on the statements of one of its speakers.¹⁵ Nor (Conclusion 12) may the author necessarily be aware of his/her idiomatic usages.

So we can't work out whether any given sentence contains an idiom which we would otherwise mistranslate literally.

Conclusion 13: *Many serious mistranslations of dead languages are undetectable in principle.*

The argument for Conclusion 11 shows that idioms quite naturally pervade languages, and the case of Amy and Barry discussing 'knowledge' shows that subtle misinterpretations which normally don't show up (of which there must be many in translations) can pop up unexpectedly and create spurious disagreements which can't be resolved. So this result follows even if we had read all the books that were ever written in that language. Furthermore, these are only the problems caused by idioms; the problems of malapropisms (below) also apply.

For translating living languages whose idioms are not all known, the same problems exist as for dead ones except that performing a sub-verbal reduction on the statements of a speaker will reveal idioms, albeit laboriously.

Communicating in living languages

Conclusion 2 tells us that knowing a language and knowing that someone is speaking it doesn't fix what their statements mean. The results for names, natural kinds, abstract terms, non-reductive rewordings and the modified Principle of Charity apply as for dead languages. So...

Conclusion 14: *We can only detect that an exchange between speakers of the same language is spurious if it is not abstract and doesn't involve names or natural kinds, and we perform reduction.*

Disagreements between speakers tend to motivate a stream of exchanges. Good arguers tend to reduce the level of discourse, and so they fairly rapidly discover either who's wrong or else that a misunderstanding is at work – even if, as with the 'vermeil' case, this is often wrongly regarded as a fallacy. Bad arguers go round in circles or stay at the same level, which is not recommended by Conclusion 10.

However, agreements are invariably static. Imagine how it would be if agreements caused conversations to continue in the way that disagreements do:

Barry: "This grass is green."

Charles: "So it is."

Barry: "But it's *green*, you fool!"

Charles: "Damn your eyes, Sir, can't you see that this grass is *green*?"

and so on. There is no possibility of discovering whether or not the exchange is spurious this way. The agreement/disagreement contrast should be seen not so much as one between confirmation and refutation, but rather between fruitless and fruitful lines of inquiry. Disagreements force us to get clearer both about what we mean, and about the facts, but agreements lull us into a false sense of security and inhibit us from getting at the truth.

Conclusion 15: *Agreements are even worse than disagreements.*

Perhaps this isn't quite fair on agreements; it would be possible – though very strange – to do a reduction process for them. However, as with disagreements, we would have to perform this reduction in almost every sentence to ensure that misunderstandings were not occurring.

¹⁵Though we *could* perform a sub-verbal reduction if we had enough contemporary captioned picture-books illustrating the 'flat earth' and so on. Perhaps the mediaeval illustrations to Mandeville's *Travels*, e.g. in British Museum Add. MS 24189, can tell us something about his (or the illustrator's) world-picture.

Tentative corollary for philosophy

Philosophical debate proceeds at a very high level, so whether or not there are different meanings in use may well not be apparent. (For instance, it may be the case that you have misunderstood this dissertation and believe some of its truths to be falsehoods, or *vice versa*.) The fact that a misunderstanding rather than an argument is going on can, however, be brought to light in at least some cases; at the paradigm level the speakers cannot be mistaken about the truth-value of what they are saying. So if we reduce to a paradigm and there is still a disagreement, then it shows that the truth-conditions of the participants' sentences are different – i.e. they are misunderstanding each other.

After-images are at the paradigm level – what could be less mistakable than something sitting right in the visual field? Consider, then, the following disagreement:

Sense-datum theorist (after looking at a bright light): “I see an after-image.”

Direct realist: “No, you don't.”

We can run through just the same argument as with Charles and his ‘vermeil’ grass; this exchange shows that this is just a misunderstanding – the sense-datum theorist and the direct realist are both right, but mean different things by ‘see’. The former means something like ‘perceive in my visual field’, the latter means ‘observe in the world’. Everyday language seems to admit *both* senses, so the direct realist is either wrong or not speaking standard English. (He is presumably here adopting a special philosophical dialect in which ‘see’ denotes only the kind of process or event involved in normal perception of the world.) At a higher level, however, a disagreement might occur which *appears* both genuine and relevant to the representationist debate, while in fact it rests on this misunderstanding.

Similarly, it is not at all clear to me that a sceptic who claims not to know things which seem to be paradigms (e.g. that this grass is green) is a speaker of English. We have already sketched how differences in opinion may be another form of misunderstanding; this may have some bearing on aesthetics. Relativism in morality suggests that some moral debates may also be misunderstandings. Unfortunately, these topics are more abstract than perception, and do not seem to possess clear paradigms. (Wittgenstein also believed many philosophical problems are just confusions about ambiguous words, like asking whether ‘sleep’ is an active or passive verb, or whether wine can be ‘dry’.)

The tentative corollary for philosophy, then, is that many unresolved philosophical disagreements may be entirely spurious; however, the difficulty or impossibility of discovering this prevents us from ever noticing.

Benjamin Finn

25th March 1991

[Submitted as BA Philosophy degree dissertation, King's College, Cambridge University, 1991. Minor revisions made October 2019. Email: ben@finn.com]

BIBLIOGRAPHY

Austin, J.L.: *How to do things with words* (Oxford: OUP, 1975)

Cupitt, D.: *The nature of man* (London: Sheldon Press, 1979)

Kripke, S.: *Naming and necessity* (Oxford: Blackwell, 1980)

Wittgenstein on rules and private languages (Oxford: Blackwell, 1982)

Mandeville, J.: *The Travels of Sir John Mandeville* (Harmondsworth: Penguin, 1983)

Quine, W.V.O.: *Word and Object* (Cambridge, Mass., 1960)