

Wittgenstein, Qualia and the Autonomy of Grammar¹
Julian Kiverstein

Forthcoming in D. Levy & E. Zamuner (Ed.'s) *Wittgenstein's Enduring Arguments* (Routledge)

Wittgenstein is commonly read as an ally of those who are sceptical about the existence of qualia.² A sceptic about qualia either flatly denies the existence of such properties or he says that supposing such properties were to exist, we could know nothing about them.³ Wittgenstein might be thought a sceptic in the first sense when he argues against the possibility of widespread colour inversion, such that we could not know that another person doesn't see colours entirely differently from ourselves. Anyone who accepts the existence of qualia will typically also embrace the possibility of such a scenario. Indeed thought experiments involving widespread spectrum inversion are standard ways of arguing for the existence of qualia.⁴

I will argue that Wittgenstein provides sound arguments against widespread spectrum inversion. However one such argument depends on a controversial assumption I will call "the effability requirement" which anyone inclined to accept the existence of qualia would simply reject. If Wittgenstein's scepticism about qualia rested on such an assumption, it would carry little weight. I will outline an alternative Wittgensteinian argument that doesn't rest on the effability requirement. However again we will see that such an argument fails to establish a more general scepticism about qualia. We must look elsewhere than to his discussion of spectrum inversion if we are to find compelling Wittgensteinian reasons for denying the existence of qualia.

Scepticism about qualia in the second sense might be attributed to Wittgenstein on the basis of his attack on the possibility of a private language. It is natural to think of qualia as properties that are known only to a single subject – the individual that has the experience to which the qualia belong. Only I can know the qualia of my experience

¹ This work was carried out as part of the European Science Foundation funded project CONTACT – Consciousness in Interaction. My thanks to the Editors for their many helpful comments.

² I shall use the term "qualia" to denote non-representational, intrinsic properties of experience, known through introspection. It is in virtue of an experience's qualia that there is something it is like for a subject to undergo an experience. Occasionally I will use the term "phenomenal character" rather than "qualia" to refer to properties of experience in virtue of which there is something it is like for a subject to undergo an experience. I will take "phenomenal character" to be a philosophically less loaded term than "qualia". Many philosophers deny that experiences have non-representational, intrinsic properties, and I take the later Wittgenstein to be amongst them, but nevertheless accept that experiences have phenomenal character.

³ Dennett (1991) – perhaps the leading exponent of scepticism about qualia – claims to be following Wittgenstein: 'Several philosophers have seen what I am doing as a kind of redoing of Wittgenstein's attack on the "objects" of conscious experience. Indeed it is....My debt to Wittgenstein is large and longstanding.' (p.462) Other philosophers who have taken Wittgenstein to be a sceptic about qualia, include Shoemaker (1982); Robinson (1994, ch.4); Kirk (1994: ch.2); Stalnaker (2000) and Block (unpublished).

⁴ See for instance Block (1990); Chalmers (1996, pp.263-66); Kirk (1994, ch.1) or Byrne (2006).

through introspection, for instance. Everyone else can know about the qualia of my experiences only by means of inference from my behaviour, and it might be objected that my behaviour isn't a reliable indicator of what I experience. Hence only I can know the qualia of my experiences – my experiences have properties that are known only to me, and are in principle unknowable to others. I will call qualia, *logically private* properties for this reason. Now Wittgenstein's private language argument has been read as an attack on the idea that our experiences have properties in virtue of which they are logically private.⁵ Wittgenstein is read as arguing that if there were a private component to experience, we would not be able to tell when this aspect of our experiences recurred.⁶ There would be no criterion of correctness entitling us to say that the item we are now identifying is the same item we encountered on an earlier occasion. But in the absence of such criteria of correctness we could not be said to know we are having an experience, and this is of course to endorse the second strand of scepticism about qualia. It is to say that even if qualia exist, we could know nothing about them.

I will argue that, appearances notwithstanding, the private language argument doesn't provide grounds for scepticism about qualia of the second kind. At best it establishes that there couldn't be a private language – a language for talking about sensations and experience understood by a single person, the user of this language. This is to say that the language we use to communicate about our experiences must be a public language. However it doesn't follow that what these elements of language refer to cannot be something private, an episode knowable in a first-person way only to the subject undergoing the episode, and in principle unknowable to others. We must therefore continue the search for an explanation of Wittgenstein's scepticism about qualia.

It is undeniably the case that in his mature writings Wittgenstein held a sceptical position about the existence of qualia, but it would be a mistake to suppose he always held such a position. In a neglected period of his writings, the so-called "middle period", we find a philosopher wholly committed to the existence of qualia. Wittgenstein described his work of this period as "phenomenology". His phenomenological project had as its aim the construction of a language for 'the description of immediate sense perception without any hypothetical addition.' (BT, §101, 349e) When we are describing immediate experience, he explains, we do not need to decide 'about the presence or absence of an object' (BT, §94, 320e). Wittgenstein goes on to write that if we try to decide whether what we are seeing is a physical object, we make the mistake of erroneously applying 'our physical mode of expression to sense data. "Objects", i.e. things, bodies in the space of the room – and "objects" in one's visual field; the shadow

⁵ For an example of this kind of reading see Robinson (1994: ch.4).

⁶ Appeal might be made to PI §258 here to support such a line of argument (see for instance, Robinson 1994: p.95-8).

of a body on the wall as an object.’ (*op cit*)⁷ I take it that such an account of sense-perception which distinguishes physical objects from sense data commits the Wittgenstein of this period to a principle Howard Robinson has dubbed the “phenomenal principle”:

“If there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that sensible quality.” (Robinson, 1994: 32)

Anyone who endorses the phenomenal principle must also accept a commitment to the existence of qualia. Take the case where there appears to be something possessing a particular sensible quality q , but where nothing exists that is q . According to the phenomenal principle one is nevertheless aware of something possessing q . What one is aware of in this case is a sense-datum and the sensible quality q is a quale, a property of this sense-datum. The middle-period Wittgenstein seems to have been committed to the phenomenal principle, but we have just seen how the phenomenal principle entails a commitment to qualia. Thus we can conclude that qualia occupied a central place in the middle-period Wittgenstein’s account of perceptual experience.

Soon after Wittgenstein undertook the project of constructing a phenomenological language, he began to express doubts about the need for such a project. In the opening section of his *Philosophical Remarks*, he wrote of phenomenological language: ‘I no longer hold it to be necessary’. I will argue that a correct understanding of Wittgenstein’s scepticism about qualia lies not in his discussion of inverted spectrum thought experiments or even in the private language argument. Instead we must understand his scepticism about qualia in the context of the considerations that led him to abandon the project of constructing a phenomenological language.

We will see how Wittgenstein gave up on the latter project when he abandoned a view of language as having a single logical structure disclosed through philosophical analysis. With this change of mind came a rejection of a theory of meaning for experience and sensation terms according to which these terms derive their significance from their referents. Instead Wittgenstein was to argue that these terms have the meaning they do because of the place they occupy in a language game. Just as our language games are not accountable to reality (PG, §68, §133; Z, §320, §331; PI, §497), so we can also say in the specific case of our sensation and experience terms that they are not accountable to experience for their meaning. It is this latter claim about the autonomy of grammar that motivates Wittgenstein’s scepticism about qualia.

⁷ It should be noted that these passages are taken from discussions in which Wittgenstein is attacking the idea of a phenomenological language. I have chosen these passages because they state particularly clearly the distinction between a phenomenological and ordinary physical language Wittgenstein will eventually reject.

1. Wittgenstein on the Possibility of Inverted Spectra

In his posthumously published *Notes for Lectures on Private Experience and Sense Data* (NFL)⁸ written between 1934-36, Wittgenstein offers some much-discussed comments on the possibility of inverted colour experience.⁹ Wittgenstein discusses two kinds of spectrum inversion, only one of which he regards as logically possible (i.e. conceivable). In the first scenario:

‘Someone says “I can’t understand it, I see everything red blue today and vice versa.” We answer “it must look queer!” He says it does and, e.g. goes on to say how cold the glowing coal looks and how warm the clear (blue) sky. I think we should under these and similar circumstances be inclined to say that he saw red what we saw blue. And again we should say that we know that he means by the words ‘blue’ and ‘red’ what we do as he has always used them as we do.’ (NFL: 284)

Ned Block (forthcoming) argues that Wittgenstein is willing to allow the possibility of this kind of spectrum inversion because a subject would be able to communicate to us how his experience had changed. The differences between the experiences of our imagined subject and normal colour experience are tracked by what the subject says about his experience. These differences can be tracked because our inverted subject uses colour terms just as we do. Hence the descriptions he gives of his experiences are ones we can understand.

Call this first kind of spectrum inversion “detectable spectrum inversion” since the subject can make known to others how his experience has changed. There is a second kind of spectrum inversion, which I will call “widespread spectrum inversion”, that Wittgenstein holds to be logically impossible (i.e. inconceivable or incoherent):

‘We said that there were cases in which we should say that the person sees green what I see red. Now the question suggests itself: if this can be so at all, why should it not always be the case? It seems, if once we have admitted that it can happen under peculiar circumstances, that it may always happen. But then it is clear that the very idea of seeing red loses its use if we can never know that the other does not see something utterly different. So what are we to do: Are we to say that this can only happen in a limited number of cases?’ (NFL: 316)

The case Wittgenstein describes in this passage is not “peculiar” but is instead something that is widespread, and so commonplace. The subjects in question have not undergone a sudden change. Perhaps they were born with crossed neural wiring so that they have

⁸ Rush Rhees published an edited version of NFL in *Philosophical Review* 1968. An unedited version can be found in Wittgenstein (1993) pp.200-88

⁹ They form the starting point for Shoemaker’s (1982) classic discussion of Inverted Spectra. More recently Block (forthcoming) has returned to these passages to mount an attack on Wittgenstein from the perspective of a friend of qualia. The discussion of Wittgenstein’s treatment of Inverted Spectra in this section is indebted to Block (forthcoming).

always enjoyed inverted colour experiences (Block (forthcoming) imagines such a scenario). They have grown up calling things red which look to them to be green and vice versa. It is this possibility Wittgenstein seems unwilling to admit.

Wittgenstein's worry seems to be that if spectrum inversion were widespread, the terms we use in reporting how we experience colour would, as he puts it, 'lose their use'. In the first kind of case we were at least agreed on the meaning of these terms, and this is what allowed us to understand what had happened to the colour experiences of the inverted subject. Not so in the case of widespread colour inversion. If spectrum inversion were widespread, there would be no agreement about how to use colour terms in describing one's experience. Some subjects will use the expression "looks red" when experiencing things that look green, others will use this expression when experiencing things that do indeed look red to them. Who is to say which of the groups is using this expression correctly? Whatever grounds we have for judging one use to be correct, these grounds would seem to count equally in favour of the other usage. Without any agreement on which usages count as normal, talk of how things look with respect to colour will be meaningless.

Of course our imagined community could use some objective feature to secure agreement about how to use expressions such as "looks red". They could agree to use this expression to refer to whatever experience a subject is having when seeing something red. The members of such a community would still lack a means of expressing what it is like for them to experience colour. There would be differences in the ways a coloured thing looked which would not show up in what subjects say about their experience. An inverted subject would say they were experiencing something that looks red when in fact they were seeing something that looks the way green things look to us.

Let us call a position like Wittgenstein's, which accepts the possibility of detectable spectrum inversion whilst denying widespread spectrum inversion, *Restricted Spectrum Inversion*. Wittgenstein's defence of restricted spectrum inversion would seem to depend on an assumption I shall call the *Effability Requirement*. The Effability Requirement says:

(ER) There are no differences in the phenomenal character of experience, which cannot be given linguistic expression.

If we assume ER, a rejection of widespread spectrum inversion follows. We have just seen how, if widespread spectrum inversion held, there would be ways that coloured things look which couldn't be given linguistic expression. Inverted subjects couldn't point to anything public to explain the different ways coloured things look to them. Nor could they use any expression of the form "looks F" where "F" stands for some colour term.¹⁰ We have just seen how an inverted subject could correctly use an expression of

¹⁰ For further discussion of this point, see section 4 of Block (forthcoming).

this form even though the way he experiences Fness is the way we experience Gness. We must conclude then that in a situation of widespread spectrum inversion there would be differences in phenomenal character that couldn't be given linguistic expression.

Unfortunately an argument against widespread spectrum inversion based on ER will fail to establish grounds for scepticism about qualia. Anyone accepting the existence of qualia will take the possibility of widespread spectrum inversion to be grounds for rejecting ER. To reject ER is to accept the existence of qualia, as has been persuasively argued by Block (forthcoming). If we are to motivate scepticism about qualia, we must either provide some additional defence of ER or look for a defence of restricted spectrum inversion independent of ER.

2. In Defence of Restricted Spectrum Inversion

Immediately following his discussion of widespread spectrum inversion, Wittgenstein goes on to add: 'we introduced the expression that A sees something else than B and we mustn't forget that this had use only under the circumstances under which we introduced it' (NFL, p.316). I take Wittgenstein to be reminding us in this passage that we have a linguistic framework for *comparing* our colour experiences with the experiences of others. It is by employing this framework that we can make sense of two individuals, A and B, experiencing one and the same property (e.g. redness) differently. In a situation of widespread spectrum inversion, no such framework could get off the ground. The expressions such a community used to talk about the ways coloured things look would cut across differences in the ways subject's experienced colours in such a way as to make these differences inexpressible. It follows that when considering such a situation we are using the expression "A experiences something else than B" outside of the circumstances or context in which it is normally used.

Detectable spectrum inversion strikes Wittgenstein as being possible because in this situation we can draw on a framework for comparing experiences. When we think about a situation in which there is widespread spectrum inversion we are considering a situation in which no such framework could get off the ground. Subjects couldn't express the way coloured things look to them in such a situation. If they cannot express the ways in which coloured things look to them, nor can they compare their respective experiences. If they have no means of comparing their experiences, how can one subject 'A' make sense of the possibility that he might be experiencing something different from another subject 'B'? Yet it is the possibility of A experiencing something different from B that we are trying to conceive of in entertaining the possibility of widespread spectrum inversion.

So it looks like Wittgenstein's defence of restricted spectrum inversion lies in considerations about what must be possible if we are to compare experiences – a

possibility that the very intelligibility of any kind of spectrum inversion may rest on. A situation in which widespread spectrum inversion held would be a situation in which subjects could make no sense of comparing their colour experiences with respect to their sameness or difference. The idea of widespread spectrum inversion presupposes the possibility of comparing colour experiences. Yet we have seen that if widespread spectrum inversion obtained, there would be no framework for comparing experiences. Thus the idea of widespread spectrum inversion is self-undermining or incoherent.

Compare the case of widespread spectrum inversion with what Wittgenstein says about pain experiences which are never displayed:

‘And remember that we admit that the other may have pain without showing it! So if this is conceivable, why not that he never shows pain; and why not that everybody has pain constantly without showing it; or that even things have pain?!’ What strikes us is that there seem to be a few useful applications of the other person’s having pain without showing it, and a vast number of useless applications, applications which look as though they were no applications at all.’ (NFL: 317)

We can make sense of someone’s being in pain but not displaying it in certain circumstances. Perhaps the person is an athlete wholly absorbed in trying to win a competition so that she is not attentively aware of her pain experience. Wittgenstein points out however that there are also a ‘vast number of useless applications’, and I take it that widespread spectrum inversion and pain experiences that are never expressed count as paradigm examples of “useless applications”. Application of the concept of pain in a situation where pain is never displayed would be a “useless” application partly because the concept of pain would have no use in such a situation. We use the concept of pain, in part, to talk about behaviour expressive of pain. If pain were never displayed, our concept of pain would lack this use. Of course this is not the only use our concept of pain is put to – in addition, we use this concept to refer to that unpleasant state we are in when some part of our body hurts.¹¹ Still a Wittgensteinian could argue that in the absence of displays of pain such applications of the concept of pain could not get off the ground. To have mastered the concept of pain is to know not just when it is appropriate to ascribe pain to oneself, but also when it is appropriate to ascribe pain to others. In a situation in which pain were never displayed, one would not know when it was appropriate to ascribe pain to others. To the extent that our concept of pain is used not just to talk about our own pain but also about the pain of others, our concept of pain would be without use.

¹¹ I do not mean to imply that this exhausts the ways in which the concept of pain is used. In addition it is used in explaining or justifying behaviour, or as a motivation or threat in persuading either oneself or another to act to mention just of the few of the many other ways this concept can be employed. My thanks to David Levy for insisting on this point.

The concept of pain would be “useless” in a second sense in such a situation. There would be no *need* for such a concept, which is to say that the concept would serve no function or purpose. We need such a concept to talk about behaviour (our own and others) expressive of pain. We use this concept to explain to others what we are feeling when we are in pain. If no one ever expressed the pain they felt, there would be no need of such a concept.

The same moral applies to the concepts we use to talk about the ways in which coloured objects look to us. In a situation of widespread spectrum inversion these concepts would have no use in either of the senses just rehearsed. There would be no need to talk about the ways coloured things look since there would be no practice of comparing experiences. Indeed if the above argument is correct, not only is it the case that these concepts *would* have no use, nor *could* they have a use. In a situation of widespread spectrum inversion there would be no agreement in how colour words are used in contexts where subjects are reporting on how colours look to them. Without such agreement the framework would be missing required for making sense of our talk about colour.

Of course it might still be objected that a practice of comparing experiences could get off the ground in a situation of widespread spectrum inversion. An individual subject that has undergone a procedure for inverting his experience could use his memory to compare his post-inversion experience with his pre-inversion experience. He could arrive at judgements about similarity or difference in experience, relying only on his memory.¹²

In response it might be argued that memory cannot provide the kind of evidence needed to warrant judgements of this kind about similarity or difference in experience. We might insist that memory could never justify a subject’s claim to know that the colours he experiences now are the same or different from the colours he experienced in the past. However, if such an argument can be made to work against the individual, it would also tell equally against a public practice of comparing experiences. I suggested above that a practice of comparing experiences can get off the ground because we are in agreement about what we mean by the terms we use to describe our experiences. However the sceptic could reasonably wonder how we know that our current use of terms is in agreement with our past usage of those terms. He could raise the very same objection it is suggested we level against the appeal to memory. On the other hand, if we suppose that there is some solution to this sceptical problem at the level of public practice, this solution could surely also be brought to bear by the individual.¹³

A more persuasive response would insist that memory can only do its work because there are certain practices already in place, in particular practices for reporting

¹² A possibility along these lines is discussed in Block (forthcoming).

¹³ See Blackburn (1984) for an objection along these lines to Kripke’s sceptical reading of Wittgenstein’s account of rule-following.

colour experiences and comparing them. Individuals can use their memory to compare earlier and later experiences but this use of memory presupposes existing practices such as the practice of giving a name to one's experience. Wittgenstein points out in *NFL*: "To give a sensation a name" means nothing unless I know already in what sort of game this name is to be used.' (1968: 290; also see *PI*, §257.) It is only because one is already a participant in a practice of naming one's sensations and experiences that one can use memory to compare earlier and current sensations or experiences.

I will have more to say in defence of such a line of argument below. However what has been said suffices for a Wittgensteinian defence of restricted spectrum inversion. Widespread spectrum inversion can be rejected because it is incompatible with our practice of comparing experiences, and without this practice we can make no sense of the idea of one subject experiencing something different from another subject.

What is the place of the Effability Requirement (ER) (the claim that there can be no differences in phenomenal character that cannot be given linguistic expression) in this defence? I've argued that the practice of comparing experiences depends on subjects being able to express the ways in which they experience things. From this claim, it follows that we cannot *make sense* of differences in phenomenal character that do not show up in what a subject says and does. This is to make a claim about our *concepts* of phenomenal character: namely that there is a constitutive connection between mastery of these concepts and what a subject says and does.

I take Wittgenstein to be committed to a semantic thesis to the effect that our sensation and experience concepts must have a public sense. I take this to be a semantic thesis, concerned with how it is possible for us talk meaningfully about our experience and their phenomenal character. It consists in part in the negative claim that we cannot talk meaningfully about our experiences using a private language – a language understood only by a single user. This is the moral of the so-called private language argument. In addition, Wittgenstein makes a positive claim to the effect that it is expressive behaviour – facial expressions, gestures and other bodily movements – which get our talk of experiences, sensations, and the like off the ground. In *PI*, §244, for instance, Wittgenstein asks how it is possible for words to refer to sensations. He returns the following answer:

'Here is one possibility: words are connected with the primitive, the natural expressions of the sensation and used in their place. A child has hurt himself and he cries; and then the adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain behaviour.'

I take “natural expressions” in this quote to refer to the facial expressions, gestures, bodily movements, exclamations *etc.* a child makes when she is undergoing a painful experience. Thus the hypothesis Wittgenstein is putting forward in this section suggests that we can use words to refer to sensations or experiences because we already have a range of expressive behaviours for communicating them. Sensation and experience terms derive their meaning from pre-established practices.¹⁴ In the absence of expressive behaviour, no context could get established for talking about our sensations and experiences. (Recall the discussion above of the possibility of a situation in which pain experiences were never displayed.) Without such a context our sensation and experience talk would be meaningless.

Thus a reasonable case can be made for the claim that we cannot *make sense* of differences in phenomenal character except by reference to expressive behaviour. However ER makes an altogether stronger claim that there are no differences in phenomenal character that cannot be given linguistic expression. The case that has just been made for rejecting widespread spectrum inversion doesn’t seem to warrant this conclusion. A proponent of qualia might concede that the differences in phenomenal character we can make sense of (i.e. conceptualise), we can make sense of only by reference to expressive behaviour. Still he will deny that all differences in phenomenal character are such that we can make sense of them. He will continue to insist that there are differences in phenomenal character, which aren’t expressible. One might wonder whether there can be differences in phenomenal character we cannot make sense of, but it is precisely such a possibility to which a proponent of the ineffable is committed. We must conclude that a Wittgensteinian defence of widespread spectrum inversion may be independent of ER. The search must continue for a Wittgensteinian defence of scepticism about qualia.

3. The Distinction between a Phenomenological Language and Ordinary Language

Perhaps a better place to look if we are to understand Wittgenstein’s scepticism about qualia is to a change that took place in his thinking in the so-called “middle period”. The Wittgenstein of this period was, for a time at least, very much committed to the existence of qualia. I will argue that the change that took place in his thinking had as a consequence a rejection of qualia. If I am right, it is these often-neglected writings that hold the key to a Wittgensteinian defence of scepticism about qualia.

Consider the distinction, which looms large in his writings from the early 1930’s, between the ordinary language used to talk about what J.L. Austin called “medium sized dry goods” (henceforth “ordinary physical language”) and what Wittgenstein sometimes

¹⁴ For further discussion of this point to which I am indebted see Hintikka and Hintikka (1986, ch.’s 8-11).

called a “primary language” or “a phenomenological language”. In *Philosophical Remarks* for instance he writes:

‘The worst philosophical errors always arise when we try to apply our ordinary – physical – language in the area of the immediately given.

If, for instance, you ask, ‘Does the box still exist when I am not looking at it?’, the only right answer would be ‘Of course, unless someone has taken it away or destroyed it’. Naturally a philosopher would be dissatisfied with this answer, but it would quite rightly reduce his way of formulating the question *ad absurdum*.

All our forms of speech are taken from ordinary, physical language and cannot be used in epistemology or phenomenology without casting a distorting light on their objects.’ (PR, §57: 88)

Wittgenstein is pointing out in this passage that sceptical problems concerning the existence of things unperceived are the result of confusing the language we use for talking about experience with ordinary language we use to talk about physical things. If we keep to ordinary physical language, the question the sceptic raises loses its force. We can give Wittgenstein’s disarming response that of course the box exists unless someone has removed or destroyed it. True, the box as-it-is-perceived-by-us doesn’t continue to exist unperceived but, as Wittgenstein notes, it is not this box whose existence the sceptic is questioning. The box the sceptic is interested in is the box we talk about using what Wittgenstein calls physical language. The sceptical problem is only pressing because we confuse these two languages.

The argument Wittgenstein employs in this passage for disarming scepticism rests on a distinction between a language which is employed for the description of immediate experience (henceforth “phenomenological language”) and ordinary physical language. When we judge that something is the case we presumably do so using ordinary physical language. We judge that the box continues to exist unless it has been removed or destroyed. Wittgenstein, at least at this stage in his writings, believed in the possibility of constructing a phenomenological language distinct from ordinary physical language

Wittgenstein was however soon to abandon this distinction between phenomenological language and ordinary physical language. In a 1929 conversation with Moritz Schlick recorded by Friedrich Waismann, Wittgenstein is recorded as saying:

‘I used to believe that there was the everyday language that we all usually spoke and a primary language that expressed what we really knew, namely phenomena. I also spoke of a first system and a second system. Now I wish to explain why I do not adhere to that conception any more.

I think that essentially we have only one language, and that is our everyday language. We need not invent a new language or construct a new symbolism, but our everyday language already is the language, provided we rid it of the obscurities that lie hidden in it.’ (Waismann, 1979: 45)

We will see how Wittgenstein conceived of a phenomenological language as uncovering the logical structure of phenomena, something he took to be covered over by ordinary physical language. He abandoned the distinction between phenomenological and ordinary physical language when he ceased believing that language has a single logical structure disclosed through logical analysis. In the next section I will argue that his scepticism about qualia may be understood as a consequence of this development in his thought.

4. Phenomenology as Grammar

Wittgenstein introduced the distinction between phenomenological language and ordinary physical language as a response to a difficulty he had identified with the conception of philosophical analysis described in the *Tractatus* (TLP). In his early writings Wittgenstein took the task of philosophical analysis to be to lay bare the logical structure of our language. A complete analysis of a proposition will clearly display this proposition's logical structure. It will reveal a set of elementary propositions connected by logical connectives out of which a complex proposition is formed. Each elementary proposition is made up of well-formed arrangements of names standing for individual objects.¹⁵ Elementary propositions do not stand in logical relations with one another: just as the states of affairs which elementary propositions describe are logically independent of each other (i.e. externally related), so too must elementary propositions be logically independent of one another.¹⁶ Wittgenstein was soon to realise that this latter thesis could not be defended given the existence of some obvious counterexamples. It was in an attempt to bring to light the logical relations between elementary propositions that he was led to introduce the idea of a phenomenological language.

Consider for instance the proposition that a single object cannot be coloured red and green at the same time. An object's being red all over logically implies that it cannot also be green all over at one and the same time. This seems to suggest that, contrary to the claims of TLP, elementary propositions are not logically independent of one another. This must be so if an elementary proposition such as <red colour patch here> can stand in logical relation of implication to a distinct elementary proposition such as <no green colour patch here> where "here" designates the same space-time location. However, the logical relations between these two elementary propositions are concealed by the subject-

¹⁵ There is some controversy about how exactly Wittgenstein conceived of the individual objects which the constituents of elementary propositions name. Answering this question is made all the harder by Wittgenstein's failure to provide any concrete examples in the *Tractatus*. Hintikka and Hintikka (1986, ch.3) argue for a reading of simple objects as objects we are acquainted with in experience (i.e. sense data). Griffin (1964) treats simple objects as physical point masses. Others have treated Wittgenstein as an anti-realist arguing that we can only identify the referent of an elementary proposition within the context of a proposition. For a discussion of the range of different possible readings see Stern (1995: pp.62-5).

¹⁶ TLP: 2.062, 4.211, 5.134, and 5.135.

predicate form of ordinary language. As Wittgenstein was to note in his 1929 essay ‘Some Remarks on Logical Form’ (RLF) the syntax of a sentence such as “This object is simultaneously red and green all over” fails to show how it is a contradiction and hence necessarily false.¹⁷ He writes: ‘It is, of course, a deficiency of our notation that it does not prevent the formation of such nonsensical constructions, and a perfect notation will have to exclude such structures by definite rules of syntax.’ (RLF: 171)

Wittgenstein conceived of a phenomenological language as a means of making perspicuous the logical relations between elementary propositions. He conceived of the construction of this language as a kind of analysis, which would reveal the logical structure of phenomena. Wittgenstein was thereby conceding that propositions do not have a *single* logical form, the familiar logical form of the propositions of ordinary language as he had supposed in TLP. A phenomenological language will disclose ‘the forms of space and time and the whole manifold of spatial and temporal objects, as colours, sounds etc. etc. with their gradations, continuous transitions, and combinations in various propositions, all of which we cannot seize by our ordinary means of expression.’ (*ibid*, p.165) He distinguished ordinary physical language from a phenomenological language – a language for describing the phenomena of immediate experience – because he believed that the propositions of ordinary language were limited in what they could show. Studying the logical form of ordinary language sentences doesn’t show, for instance, the many different degrees of brightness we can see, or the variety of shades between red and blue. Nor does it show us the structure of the relationships between the colours we see. We need another sort of analysis to uncover relationships such as these, a task which he initially seems to have thought we could achieve through the construction of a phenomenological language.

In the *Philosophical Remarks* (PR), Wittgenstein introduces a colour octahedron to represent the full range of relations in which colours can stand in to one another. ‘The colour octahedron’ he writes ‘is grammar, since it says that you can speak of a reddish blue but not of a reddish green, etc.’ (PR, §39) A colour octahedron will clearly show which colour combinations we can meaningfully say are possible and which we cannot meaningfully talk about. Such an apparatus succeeds where ordinary physical language fails, since it constitutes a notation for correctly representing the logical structure of experience.

We now have a handle on why Wittgenstein came to distinguish phenomenological language from ordinary physical language. We saw at the end of the previous section

¹⁷ Wittgenstein wrote this essay in the summer of 1929, his first piece of philosophical writing following his abandonment of philosophy in 1918. He was apparently extremely dissatisfied with the results, refusing to read the paper at the joint session of the Aristotelian Society for which it was written. No doubt his dissatisfaction was in part due to doubts about the notion of a phenomenological language he employs in this essay.

how this distinction was one that Wittgenstein soon came to repudiate. In the opening section of the *Philosophical Remarks* for instance, Wittgenstein writes that he no longer has a phenomenological primary language as his goal – he says that he no longer views such a language as “necessary” (PR, §1). To give up the goal of a phenomenological language is to give up the task of constructing a notation capable of representing the structure of experience. So what led Wittgenstein to this change of mind?

I will follow Stern’s (1995) suggestion that Wittgenstein’s abandoned his goal of a phenomenological language because he abandoned the idea of there being a ‘single unitary order that was concealed by our ordinary language...a general form of the proposition shown by the structure of language as a whole.’ (p.138)¹⁸ Stern describes how over the course of the 1930s Wittgenstein came instead to view language as ‘a family of interrelated games...a number of autonomous systems of representational conventions, each with its own logical structure.’ (*ibid*) Wittgenstein rejected the goal of a phenomenological language because he rejected the project he had pursued in TLP, of constructing a form of representation which would make perspicuous what is only confusedly expressed in ordinary language. He came to the view that propositions do not have a single general form which logical analysis can uncover, a view we already find expressed in RLF where he introduced the idea of a phenomenological language. There is no single notation that can make perspicuous what is concealed by ordinary language. Instead there are a multitude of different methods of representation or notations each of which serves a different purpose.

Once he abandoned this view of language (as having a single logical structure), he no longer needed to think of phenomenological language as having a life distinct from that of ordinary physical language: ‘Our ordinary language is completely in order, as long as we are clear about what it symbolises.’ (Waismann, 1979: 46) We get clear about what ordinary language symbolises by examining and comparing different methods of representation or notations so as to lead us to a recognition of ‘what is essential and what inessential in our language if it is to represent, a recognition of what parts of our language are wheels turning idly.’ (PR, §1) Such a project he goes on to say ‘amounts to the construction of a phenomenological language.’ (PR, §51)

It took Wittgenstein a short time to fully recognise the implications of his rejection of the necessity of formulating a phenomenological language. If a phenomenological language is not distinct from ordinary language, the sentences we use to talk about our experience must get their meaning in the same way as any other sentence of ordinary language. Yet throughout 1929 he continues to write of propositions describing experience as deriving their significance from experience:

¹⁸ Noë (1994: pp.15-20) defends a similar reading.

‘The phenomenon is not a symptom of something else; rather it is reality.

The phenomenon is not a symptom of something else which then makes the sentence true or false; rather, it is itself that which verifies it [the sentence].’ (MS107, pp.223-4, quoted by Stern (1997: p.139)

By the 1930’s Wittgenstein rejected this account of meaning. He came to conceive of phenomenology as concerned with, to borrow Noë’s (1994) expression, ‘what it makes sense to say’ about experience. The rules which determine what it makes sense to say are rules of grammar. Moreover it is rules of grammar which determine ‘what is essential’ (PR §1) and what is not in our language. Grammar determines which parts of our language ‘are wheels turning idly’ (*ibid*). The grammar of our talk about colour experience determines, for instance, that it makes sense to speak of reddish-yellow but not of reddish-green or that there can only be one colour at one place at any given time:

“‘Red and green don’t go together at the same place’ does not mean, they are never actually together, rather it means that it is nonsense to say that they are at the same place at the same time and therefore also nonsense to say that they are never at the same place.’ (BT, §100: 342e)

Wittgenstein came to conceive of a word’s *use* as determining its meaning. In §19 of BT (pp.63e-64e), for instance, Wittgenstein asks why the proposition ‘one can’t divide red’ makes no sense. The answer he returns is that this sentence is without meaning because ‘it doesn’t belong to the particular game that, judging from its appearance, it seems to belong to.’ Games are of course defined by the rules which determine, for instance, which moves within the game are legitimate, and which illegitimate. So Wittgenstein is saying that the rules of the game, of which our talk of colours is a part, determine that the sentence ‘one cannot divide red’ is without meaning. It is these rules which fix what it makes sense to say about colour experience.

Consider once again the proposition: “two primary colours cannot simultaneously occupy the same place at the same time” (call this the “colour exclusion proposition”). Wittgenstein came to view the colour exclusion proposition as expressing a rule of grammar which part determines what it makes sense to say about colour experience. As an expression of a rule of grammar this proposition cannot be justified by appeal to any empirical fact – to a fact about the intrinsic character of our colour experience for instance. Wittgenstein wrote in his *Philosophical Grammar*: ‘Grammar is not accountable to any reality.’ (PG, 133) Any empirical statement we formulate relating to colour will depend for its meaning on the very same principles of grammar we are trying to justify. If we try to justify our colour grammar by appeal to the fact that “there really are four primary colours” we are presupposing our ways of dividing up the world (see Z, §331 & 357 and Forster (2004: pp.33-39)). This division of colour space is of course what we are

trying to justify. What is it then that accounts for our being warranted in asserting the proposition “two primary colours cannot be in the same place at the same time”, if it is not some empirical fact?

To answer this question it will be helpful to turn briefly to a passage from the *Investigations*:

‘So does it depend wholly on our grammar what will be called (logically) possible and what will not – i.e. what the grammar permits? – But surely that is arbitrary! Is it arbitrary? – It is not every sentence-like formation that we know how to do something with, not every technique has an employment [*Verwendung*] in our life; and when we are tempted in philosophy to count some quite useless thing as a proposition, that is often because we have not considered its application [*Anwendung*].’ (PI, §520)

What this quote suggests is that the colour exclusion proposition is a statement of grammar because this proposition has an application in our lives. Those grammatical statements that lack such an application in our lives do not genuinely express propositions, and thus do not genuinely possess a meaning even though they may seem to. They are like the useless propositions of the philosopher we encountered above. (Recall the discussion of widespread spectrum inversion and of pain that is never displayed.) Thus a proposition genuinely expresses a rule of grammar if the rule in question has a function, which is to say that the rule achieves some practical purpose in our lives. It is only because the colour exclusion proposition is useful – i.e. fulfills a practical purpose in our lives – that it counts as a rule of grammar.

What practical purpose does the colour exclusion proposition serve? Why couldn’t its contrary (i.e. the statement that two primary colours *can* occupy the same place at the same time) serve us just as well? While it is true that we cannot, without the threat of circularity, appeal to any facts to *justify* our use of such a grammatical principle, we can however appeal to facts to explain a rule of grammar’s usefulness. A rule like the colour-exclusion rule is useful because certain empirical regularities obtain. If different regularities obtained, different rules would be useful and the colour exclusion rule would be useless. In his *Remarks on the Foundations of Mathematics* (RFM), Wittgenstein writes: ‘Calculating would lose its point, if confusion supervened. Just as the use of the words “green” and “blue” would lose its point...the arithmetical proposition would not be false, but useless, if confusion supervened.’ (RFM, III, §75) Our talk of green and blue is useful only because it doesn’t depend on confusion, but is instead a way of classifying colours in accordance with the regularities we find in experience. If only confusion prevailed when we made use of such a classification, the rules which govern the use of these colour terms would be useless. We have the rules of grammar we do because they

are *useful*. In a situation where only confusion prevailed we wouldn't have the rules of grammar we do, we would employ some alternative rules.¹⁹

I have suggested that understanding why Wittgenstein abandoned the distinction between a phenomenological and an ordinary physical language will help us to understand his scepticism about qualia. We are finally in a position to draw some conclusions.

First we can understand Wittgenstein's scepticism about qualia as the result of his rejection of the distinction between phenomenological and ordinary physical language. A proponent of qualia will say that when we talk about the qualitative character of our experiences, the terms we use have their meaning partly in virtue of the qualia belonging to our experiences. Consider once again, the colour exclusion proposition: a proponent of qualia will say that what makes this proposition true is the intrinsic character of our colour experiences. We cannot experience two primary colours as simultaneously occupying the same place at the same time because an experience can only have a single intrinsic qualitative character at any given time. Just as an object cannot be in two different places at the same time so experiences cannot have two different intrinsic characters at the same time. In a similar fashion, it is the intrinsic character of our colour experiences which explains why it is possible for us to experience reddish yellow, but impossible for us to experience reddish green. If we say that the meaning a proposition expresses is fixed by a proposition's truth conditions, it follows that propositions which make reference to the qualitative character of our experiences have their meaning partly in virtue of qualia. It is the fact that an experience instantiates qualia which, according to proponents of qualia, makes a proposition concerning the qualitative character of an experience true.

This sort of explanation of colour exclusion is just what Wittgenstein came to reject once he abandoned the distinction between phenomenological and ordinary physical language. The colour exclusion proposition, and other similar examples, seemed to show that propositions describing immediate experience had a logical structure which was hidden by ordinary language. Wittgenstein came to realise that it was a mistake to

¹⁹ It might be argued that the appeal to practical purpose is incompatible with the autonomy of grammar claim. Am I not explaining our use of certain rules by reference to the fact that these rules are useful? Doesn't this amount to saying our language games are accountable to reality – to the fact that each of the moves in a language game serves some practical purpose? I don't think the appeal to practical purpose has this consequence. There is, on the one hand, the question of what justifies our employing a particular rule. In answer to this question I have made appeal to Wittgenstein's talk of rules of grammar as having an application in our lives (see passage from PI §520 quoted above). On the other hand, there is the question of why we follow one rule rather than its contrary. Wittgenstein is quite clear that our employing a particular grammatical principle isn't a matter of arbitrary choice (in addition to PI §520 see Z §358, WL, p.70 and LC, p.49 and Forster 2005: ch.3). My suggestion is that while practical purpose cannot justify our employing the grammatical principles we do, it nevertheless acts as a constraint explaining in part why our applying the rules we do isn't arbitrary. For fuller discussion to which I am indebted see Forster 2005.

think of language as having a single logical structure. With this realisation he abandoned the view that phenomena have a particular logical structure in virtue of which certain ways of talking about experience are meaningful and other ways of talking are meaningless. He came to believe that we talk about experience as we do, because of the rules of grammar which fix what it makes sense to say about experience. What makes it the case that certain ways of talking are permitted and others are proscribed is ultimately whether these ways of talking serve a practical purpose in our lives.

Wittgenstein's brief flirtation with a phenomenological language taught him that we talk about colour experience as we do not because of the facts of colour experience. We cannot appeal to our colour experiences having particular intrinsic, non-representational properties, as proponents of qualia would have us do, to explain why some ways of talking are meaningful and others meaningless. No putative factual claim such as this one can explain why we divide up colour space as we do. The meaningfulness of any such factual claim derives from our employment of certain rules of grammar. It cannot explain our employment of these rules of grammar without lapsing into vicious circularity. We have seen how proponents of qualia explain the meaning of our experience and sensation talk by appeal to our acquaintance with qualia. Wittgenstein's scepticism about qualia lies in his rejection of this account of meaning. A proponent of qualia is committed to saying that the terms we use to talk about the qualitative character of our experience have their meaning in virtue of qualia. Wittgenstein rejects this account of meaning. Hence he rejects the existence of qualia.

It might be thought that an argument can be given for the existence of qualia quite independent of a commitment to any particular theory of meaning. However this doesn't seem to be the case for some of the leading arguments given in support of qualia. Take Frank Jackson's (1982) knowledge argument. When Mary, the expert on colour vision, is released from her black and white room she gains some new propositional knowledge. A proponent of qualia will say that the propositions she learns make reference to colour qualia. He will say that the propositions Mary could not know prior to her having colour experiences are propositions which have their meaning in virtue of colour qualia. Alternatively consider Ned Block's (1990) inverted earth thought experiment, which is supposed to establish that a subject's intentional states can change their content while the phenomenal aspects of their experience remain the same. I take it that the force of this argument is supposed to be that there is something about the meaning of our colour terms which changes and something that stays the same. To explain what it is about the meaning of our colour terms that stays the same we have to appeal to qualia – non-representational, intrinsic properties of experience. So it would seem that another of the major arguments for qualia depends on considerations to do with a theory of meaning for experience terms.

According to the reading I have been advancing Wittgenstein is a sceptic about qualia because of his views about the grammar of our talk about sensations and experiences. He argues that such grammar isn't explained by any factual claims, such as the claim that our experiences have particular qualia. Rather such factual claims are made meaningful by the rules of grammar that regulate our experience and sensation talk. What explains the existence of these rules is the practical purpose they serve in our lives. It is not only proponents of qualia who claim that our experience talk acquires its meaning from acquaintance with our experiences. This is also a commitment of those who posit the existence of phenomenal concepts in accounting for the semantics of our talk about sensations and experiences. These philosophers have appealed to the existence of phenomenal concepts in arguing for the conclusion that the so-called explanatory gap (see Levine 1997) is a cognitive illusion (see Loar (1990/1997); Tye (2000); Papineau (2002) and Perry (2001)). However if the account I have given of Wittgenstein's scepticism about qualia is correct, it not only puts into question that account of the meaning of our experience and sensation terms given by proponents of qualia. It also challenges those who would defend materialism by appeal to phenomenal concepts. It is with this challenge that I will conclude.

5. Phenomenal Concepts

Proponents of phenomenal concepts typically defend what Chalmers (1996) has called "type B materialism". Consider Jackson's knowledge argument involving Mary the colour-scientist who, confined to a black and white room for her entire life, learns the complete science of colour. Type B materialists accept that Mary learns something new when she experiences colour for the first time, but they deny that she learns a new fact. Instead they say she acquires concepts that allow her to think about facts she already knew in a new way. Some type B materialists suggest we think of phenomenal concepts as recognitional concepts (e.g. Loar (1990/1997); Tye (2000); Carruthers (2000)). Phenomenal concepts have a characteristic functional role – they allow us to recognise the differences between phenomenal qualities and state directly on the basis of introspection. Since what is recognised is a physical state, the deployment of phenomenal concepts does not require Mary to think about a new object or property distinct from the ones she already thought about in mastering colour science. Her mastery of phenomenal concepts enables Mary to think about the same physical states she was thinking about before but in such a way that she can discriminate between them based on introspection.

Perry (2001) argues that we should conceive of phenomenal concepts as functioning like indexical expressions. When Mary sees red for the first time she acquires concepts that allow her to run through the following chain of thoughts:

- (1) Q_R is what it is like to see red (where “ Q_R ” is the concept she knows others employ in thinking about what it is like to see red).
- (2) *This* is what it is like to see red (the thought Mary has when attending to her first ever experience of seeing red)
- (3) Q_R is *this* subjective character

So when Mary has an experience for the first time she acquires an ability to think thoughts of the form (2) and this ability allows her to learn identity statements of the form (3). Crucially (3) doesn’t describe a new fact Mary has learned. Instead it tells us something about Mary’s ways of thinking about redness.

As a final example, consider Papineau’s (2002) quotational account of phenomenal concepts. According to Papineau, when Mary has a colour experience for the first time she acquires abilities – the ability to imaginatively re-create and introspectively classify colour experiences. These abilities enable Mary to think new thoughts, thoughts she couldn’t entertain prior to having her experiences of colour. By imaginatively re-creating her experience she can think for instance, “This is what people experience when they look at ripe tomatoes.” When Mary imaginatively recreates an experience, the concepts she employs in forming thoughts like this one literally co-opt an instantiation of the experience being thought about. On this view when we form a thought using a phenomenal concept we reproduce an example of the experience the thought is about. This exemplar is then included in the content of thought. When we deploy a phenomenal concept in an act of imagination, what this concept refers to are experiences that appropriately resemble the imaginatively recreated experience (c.f. Papineau 2002: 118-9). When we use a phenomenal concept in forming an introspective thought, what the phenomenal concept refers to will be an act of perceptual classification. Introspective uses of phenomenal concepts *include* the experiences they refer to.

What these different accounts of phenomenal concepts have in common is that they all posit concepts available only to persons who have undergone the experiences the concepts have as their referents. They deny that these concepts refer to qualia but they agree with proponents of qualia that to master a phenomenal concept, a subject must be acquainted with the experiences the concept designates. They just disagree about the nature of the experiences in question, with proponents of phenomenal concepts arguing that such experiences are either physical or functional in nature.

We have seen how proponents of qualia claim that the terms we use to talk about subjective experience derive their meaning from what they designate, the experiences themselves and their qualia. The story is slightly more complicated in the case of proponents of phenomenal concepts who tend to tell different stories about (1) what fixes a concept’s meaning, and (2) the conditions a subject must satisfy if she is to

possess a concept. They agree that a subject must be acquainted with an experience if she is to deploy a phenomenal concept in thinking about the experience in question. However they deny that it is acquaintance with an experience that fixes the meaning a concept expresses. Despite this complication, to which I will return in a moment, proponents of phenomenal concepts and qualia are ultimately agreed that the meaning a phenomenal concept expresses derives from the experience the concept has as its referent. It is this claim that we have seen Wittgenstein dispute. If my reading of Wittgenstein is correct, he would argue that we can refer to such items as sensations or experiences only because we make utterances and form thoughts according to publicly defined rules of grammar which determine what it makes sense to say in this domain. These publicly defined rules make it possible for us to compare experiences and so to identify and reidentify experiences. It is through mastery of a shared language that we can compare experiences and hence identify them.

Papineau is the only proponent of phenomenal concepts to directly engage with Wittgenstein's arguments, so far as I am aware. He writes:

‘content does not derive from normative rules, but rather from the kind of non-normative natural facts invoked by causal or teleosemantic theories of representation...So, on my view, it is no deficiency in Mary's concept that she is not sensitive to any normative principles trying its use to public criteria. It is enough that her concept has appropriate causal or teleosemantic credentials, since this in itself will ensure that her concept refers determinately, and that judgements made by using it have definite truth conditions.’ (Papineau, 2002: 129)

According to Papineau, a phenomenal concept derives its meaning from facts about the causes and biological function of uses that are made of those concepts. Phenomenal concepts, he conjectures, may have helped our ancestors to better predict the behaviour of others, to anticipate their own future experiences, to facilitate reflection on the epistemic standing of their own beliefs (2002, footnote p.116). For each phenomenal concept there will be something about the use that was made of the concept by our evolutionary ancestors which Papineau is gambling will enable us to fix its content. Thus there is no need to appeal to grammar to do the work of fixing what it makes sense to say about experience. The biological function of our phenomenal concepts will do this work.

There is a dispute between Papineau and a Wittgensteinian about how the meaning of our phenomenal concepts is fixed. I want to finish up by suggesting that it is an advantage of the Wittgensteinian account that it doesn't separate the question of how the meaning of our concepts is fixed from the question of what conditions a user of concept must satisfy if he is count as possessing a concept. Even if Papineau has a satisfactory story to tell about how phenomenal concepts get their meaning, a question on which I

will remain neutral, he doesn't have an adequate story to tell about what it takes to master a phenomenal concept.

Arguably to understand a phenomenal concept one must know which experience it is the concept refers to – one must be able to identify the experiences of the type the concept denotes. Papineau describes two forms that such an understanding can take (2002, pp.129-139). The first account he suggests requires a subject to know that people acquire phenomenal concepts concerning their experiences upon having those experiences:

‘...let us suppose that Mary has a companion, Jennifer, who similarly has never seen colours but knows all about colour vision in material terms. Jennifer isn't shown the piece of paper that Mary sees but is told in material terms that it is red. Then Jennifer who knows that people acquire phenomenal concepts of experience once they have had those experiences, will be able to understand Jane's 'qual' as expressing just such a phenomenal concept, and indeed one which refers to the experience caused in her by seeing something red.’ (*op cit* p.130)

Does Jennifer have an understanding of the phenomenal concept Mary employs in this case? She understands that the concept Mary has employed has as its referent, the experience caused in Mary upon seeing something red. It is not clear however that Jennifer knows which experience this is, because it is not clear that a description of this kind is a uniquely referring expression. From the possibility of spectrum inversion it follows that there is more than one experience Mary might be undergoing caused by her seeing something red. It follows that grasping this description isn't sufficient to give Jennifer knowledge of which experience Mary is undergoing. Yet it is plausibly a condition on grasping a phenomenal concept that Jennifer be able to identify which experience the concept has as its referent.

The second, more demanding, account Papineau describes requires Jennifer to deploy her own phenomenal concept in identifying the experience the concept denotes. The difference from the first case is that this time we allow that Jennifer can see the red piece of paper which Mary can see:

‘Then she (Jennifer) too will acquire the phenomenal concept of seeing something red, and will thus be able to think about the experience referred to be 'qual' in the same way in which Mary does, and not just materially.’
(*ibid*)

Is this really an improvement on the first account? Can Jennifer know which experience it is that Mary is referring to when she uses the expression 'qual' just by deploying her own 'qual' concept? In order for us to say that Jennifer understands Mary's 'qual' concept we must have grounds for saying that when she deploys her own 'qual' concept she is

applying the very same concept as Mary. We need to have grounds for saying that she is thinking about seeing something red in the same way as Mary. Papineau seems to assume that because both subjects use their 'qual' concepts to think about experiences they are having in the presence of something red they must be using the same concept. However, again we can say that the two subjects may be having different experiences in the presence of something red – one may see something that looks red, the other may see something that looks green. If they are having different experiences, they will each be deploying different phenomenal concepts. Again we have no grounds for saying that Jennifer has understood what Mary means when she uses the word 'qual' to express her phenomenal concept.

To be fair to Papineau he does notice this difficulty, and he rejects the possibility of spectrum inversion on *a posteriori* grounds (2002: pp.131-2 and ch.7). Even if we suppose Papineau has good reasons for rejecting spectrum inversion, still my point stands that he has not given us an adequate account of what is involved in understanding a phenomenal concept. Of course it is probably the case that Papineau does not assign much importance to a theory of understanding. He does after all take concepts to be individuated by their biological function as we have already seen. It is a concept's biological function – roughly the conditions under which it is used in such a way as to further an organism's fitness – which fixes what the concept has as its referent. I will return to this point in my conclusion.

A Wittgensteinian way around difficulties of this kind would be to invoke the metaphor of moves within language games. How does Jennifer know which experience it is that Mary is talking about? There are rules governing the use of the word 'qual' that fix how this word must be used if it is to be used correctly. This rule determines which linguistic moves Mary can legitimately make with this word. The word 'qual' is used correctly when it is used to refer to the experience Mary undergoes when she experiences something red. What rule must Jennifer follow if she is to grasp how the word 'qual' must be used?

Here it will be useful to introduce a distinction Hintikka and Hintikka (1986: ch.11) make between two kinds of language games, which they call 'primary' and 'secondary' respectively. In a primary language game a subject gives spontaneous expression to an experience through a gesture, expressive bodily movement or exclamation. Using a secondary language game we can talk about an experience separately from its behavioural manifestation as is the case when someone lies; simulates or play acts (see PI, II, §xi). However our doing so in an important sense presupposes the establishment of meaning brought about by the primary language game. Thus our use of the word 'pain' for instance, is given its primary meaning by spontaneous reactions to pain such as exclamations, bodily movements, facial expressions etc. We can subsequently replace

such spontaneous expressions of pain with the use of the word 'pain'. When we do so, we replace one means of expressing pain with another. Someone can then pretend to be in pain thus breaking the connection between pain behaviour and pain. However they can do so, only because the primary language game is already established.

We understand which experience 'pain' refers to then because of the constitutive connection between pain and expressions of pain, of which the use of the word 'pain' is one instance. Understanding the word 'pain' is thus knowing the place this word has in a language game – a game in which pain experiences are expressed by a variety of different behaviours. Can we make an analogous move for Mary's use of the word 'qual' to explain under what conditions Jennifer can be said to understand this term and so know which experience it refers to?

'Qual' is of course different from 'pain' in that it denotes a colour-experience – the experience Mary has when she sees red for the first time. What is the language game in this case which Mary's use of this word depends on for its meaning? The following passage seems to address our question directly:

'How is he to know what colour he is to pick out when he hears "red"?
– Quite simple: he is to take the colour whose image occurs to him when he hears the word. – But how is he to know which colour it is 'whose image occurs to him'? Is a further criterion needed for that? (There is indeed such a procedure as choosing the colour which occurs to one when one hears the word "...')

"Red' means the colour that occurs to me when I hear the word 'red'"
– would be a definition. Not an explanation of what it is to use a word as a sign.' (PI, §239)

The language games we play with colour are extremely complex but one facet of this game will be that our uses of colour terms tend to go together with visual impressions of certain sorts. Mary's use of 'qual' is much like an exclamation of pain that occurs along with a pain sensation. It institutes a connection between the private mental episode known in a first person way only to Mary – the visual impression she is undergoing – and an expression in a language game. This relation between the visual impression and the use of the word is a grammatical or conceptual: it fixes the meaning of the word. It is because of this connection between use of a colour term and the occurrence of some visual impression that the visual impression as something private, known in a first person way only to its subject, can enter into a language game.

Superficially this looks like the same story that proponents of phenomenal concepts tell about how we can think about our conscious experiences. On Papineau's account, for instance, when Mary thinks about her first experience of redness she does so by including the token experience she is undergoing in her thought. It looks like

Wittgenstein is proposing something similar in saying that 'red' means the colour that occurs to me when I hear the word 'red'. Isn't this equivalent to saying that the phenomenal concept 'red' refers to the colour that occurs to me when I token this concept?

There is however a key difference. On Wittgenstein's story the colour term 'red' gets its meaning from its place in a language game. Rules of grammar determine the moves that can be made within this language game that fix how this term is to be understood. It is because the visual impression of redness is correlated with a language game that we are able to make ourselves understood to one another in talking about our visual impressions. Proponents of phenomenal concepts do not have anything to play the same role as language games, to which they can appeal in explaining how we make sense of phenomenal concepts. This is why they fail to explain how Jennifer can know which experience it is that Mary picks out when she uses 'qual'. If I am right, the Wittgensteinian appeal to language games escapes this difficulty.

Now it might be objected that the Wittgensteinian hasn't explained how Jennifer can know which experience it is Mary is referring when she uses her 'qual' concept, anymore than the proponent of phenomenal concepts. All that has been said above is that 'qual' refers to a visual impression correlated with a token use of this term. However it could be that each person has a different visual impression when they use this term.

Wittgensteinians have a response to this objection: they can appeal to the same argument that was deployed against the possibility of widespread spectrum inversion. In order for us to talk meaningfully about colour experiences there has to be some agreement about how the terms constitutive of such talk are to be used. It is this agreement that makes it possible for us to compare our colour experiences. In the absence of this agreement, talk of how things look with respect to their colour would lack meaning and the practice of comparing experiences would not be able to get off the ground. Yet the very idea of differences in experience, which this objection raises, rests on the possibility of our making sense of comparing experiences. So while it is possible for us to differ in the visual impressions we undergo when experiencing colour,²⁰ the language games we engage in talking about such experiences allow us to sidestep such possibilities of variation. They establish sufficient agreement in meaning for us to talk about and compare experiences.

²⁰ Wittgenstein repeatedly concedes this much, see for instance PI, §272 where Wittgenstein writes: "The essential thing about private experience is really not that each person possesses his own exemplar, but that nobody knows whether other people also have this or something else. The assumption would thus be possible – though unverifiable – that one section of mankind had one sensation of red, another section another."

The key advantage the Wittgensteinian has over both the proponent of qualia and the proponent of phenomenal concepts is that he can explain how it is possible for a subject to refer to subjective experience. Both proponents of qualia and of phenomenal concepts lack a satisfactory answer to this question. I will conclude by making this failing explicit.

The proponent of qualia takes the meaning of our experience and sensation terms to be fixed by their referents but they fail to explain how it is possible for us to refer to our experiences and sensations. This can be brought out through reflection on the possibility of widespread spectrum inversion a commitment to qualia entails. Suppose that the meaning of our experience terms was given by their referents. Further suppose that our experiences have qualia such that I have no way of knowing that what I refer to when I talk about my experiences is not something different from what you refer to when you talk about your experience. If we concede the latter possibility, there could be no agreement in what we mean by our experience terms. If we have not accounted for such agreement, we have arguably not accounted for how it is possible to refer to colour experiences. We will not have explained how it is possible for different individuals to refer to one and the same experience.

A proponent of phenomenal concepts makes appeal to evolutionary function to explain what fixes the meaning of our experience and sensation terms, distinguishing the question of what fixes a concept's meaning from the question of what conditions must be met if a subject is to possess a concept. However because they fail to return a satisfactory answer to this latter question, they fail to explain how it is possible for us to refer to our subjective experiences. They fail to provide an account of what it is to possess mastery of a phenomenal concept because they fail to explain how the user of the concept can know which experience it is a phenomenal concept designates. This is to fail to explain how it is possible for phenomenal concepts to refer.

The Wittgensteinian doesn't separate a theory of meaning for phenomenal concepts from a theory of understanding. Thus he has a satisfying story to tell about how it is possible for us to refer to our experiences and sensations, something that we have seen proponents of phenomenal concepts and proponents of qualia both fail to explain. Some may regard the autonomy of grammar thesis too high a price to pay. However my suspicion is that it is only once we recognise the autonomy of grammar that we will find a satisfying story to tell about how it is possible for us to talk about subjective experience and make ourselves understood.