

## The importance of concepts

Article (Accepted Version)

Sawyer, Sarah (2018) The importance of concepts. *Proceedings of the Aristotelian Society*, 118 (2). pp. 127-147. ISSN 1467-9264

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/75872/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

### **Copyright and reuse:**

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

## THE IMPORTANCE OF CONCEPTS

SARAH SAWYER

This is a draft. The final paper will be published in:

*Proceedings of the Aristotelian Society* (2018) Vol CXVIII, Issue 2

Podcast and draft paper available online at:

<https://www.aristoteliansociety.org.uk/the-proceedings/the-2017-18-programme/sarah-sawyer/>

*Abstract.* Words change meaning over time. Some meaning shift is accompanied by a corresponding change in subject matter; some meaning shift is not. In this paper I argue that an account of linguistic meaning can accommodate the first kind of case, but that a theory of concepts is required to accommodate the second. Where there is stability of subject matter through linguistic change, it is concepts that provide the stability. The stability provided by concepts allows for genuine disagreement and ameliorative change in the context of conceptual engineering.

### I

*Introduction.* The meaning of a word can, and typically does, change over time. I take meaning shift as a datum. Here are some examples. The term ‘meat’ used to mean food in general, and used to have a correspondingly broader extension than it has today; ‘a clue’ used to mean a ball of yarn, and hence had an entirely different extension from the one it does today; and the term ‘spinster’ used to mean a woman who spun wool, and thus had an extension that is distinct from but, perhaps, partially overlapping with the one it has today. There are historical, etymological connections between the old and the new meanings in each of these cases. I leave at an intuitive

level the question of the individuation of words and focus instead on providing an account of meaning shift.<sup>1</sup>

From a philosophical perspective, meaning shift falls into two broad categories: cases that are accompanied by a corresponding change in subject matter; and cases that are not.<sup>2</sup> The examples just given ('meat', 'yarn', and spinster') are all of the former kind. In section II, I provide an account of linguistic meaning that accommodates such cases. The cases involve a single word  $w$  whose meaning  $m$ , extension  $e$ , and subject matter  $s$  at time  $t$ , are different from its meaning  $m'$ , extension  $e'$  and subject matter  $s'$  at a later time  $t'$ . The account of linguistic meaning I provide explains the difference between  $m$  and  $m'$ , the account of the relation between meaning and extension I provide explains the corresponding difference between  $e$  and  $e'$ . And, because the change in extension in these cases maps onto a change in subject matter over the relevant time period, there is no additional explanation required of the difference between  $s$  and  $s'$ . The account of linguistic meaning I provide is therefore fully explanatory of meaning shift in these cases.

The second kind of case, which I discuss in section III, is not so straightforward. This is because, in the second kind of case, a shift in meaning is not accompanied by a change in subject matter even though, in a sense to be made clear, the extension of the term changes as a result of the meaning shift. It is this combination of difference in meaning and sameness of subject matter that provides the theoretical complication. Examples of meaning shift of this second kind include the change in meaning of terms such as 'whale' and 'atom', which are associated with distinct theories of the same subject matter at different times, and the ameliorative

---

<sup>1</sup> There is reason to think words cannot be individuated by appeal to syntactic features given that identity of shape is neither necessary nor sufficient for identity of words. The phenomenon of meaning-shift casts doubt on the possibility of individuating words by their semantic features. For discussion of identity and persistence conditions of words see Kaplan (1990). See also Kaplan (2011), Hawthorne and LePore (2011) and Bromberger (2011).

<sup>2</sup> In this paper I deal with general terms only and leave indexical terms and singular uses of proper names to one side.

change in meaning of terms such as ‘consent’ and ‘rape’, which, I argue, are also underwritten by different understandings of the same subject matter at different times.

In cases of this second kind, there are two distinct phenomena that are in need of explanation—difference in meaning on the one hand, and sameness of subject matter on the other. This amounts to a complication because no single element can explain both. Consequently, this second kind of meaning shift cannot be adequately explained by an account of linguistic meaning alone. While my account of linguistic meaning given in section II provides an explanation of the first phenomenon of difference in meaning, I argue that an account of concepts is required to provide an explanation of the second phenomenon of sameness of subject matter. In section IV, I present an account of concepts that, in conjunction with the earlier account of linguistic meaning, accommodates cases of meaning-shift of this second kind.<sup>3</sup>

The dual appeal to a theory of linguistic meaning on the one hand and a theory of concepts on the other is underwritten by a substantive distinction between language and thought. Words, I maintain, are two-sided: they *have* linguistic meanings, and they *express* concepts. Where there is stability of subject matter through linguistic change, I argue, it is concepts that explain the stability. The linguistic meaning of a term cannot be equated with the concept expressed by that term given that the two have different explanatory roles and are subject to different norms.

The distinction between the linguistic meaning of a term and the concept expressed by that term provides an understanding of genuine disagreement, thus avoiding widespread linguistic and conceptual relativism, and it provides a framework for understanding ameliorative change in the context of what has become known as conceptual engineering. These implications are discussed in section V. The claims of the paper are general, applying to mathematical and logical terms, to theoretical terms in the natural, social and political sciences, to philosophical terms and to overtly normative terms such as ethical terms. Unfortunately, I do not have the space to work

---

<sup>3</sup> One might say, as Rob Stainton suggested in discussion, that cases of both kinds involve the preservation of a phonological item, but only cases of the second kind involve the preservation of a lexical item. The distinction between the categories is inevitably vague given the distinction, which I articulate below, between subject matter and extension. Nonetheless, there are determinate cases in each category, and hence a fundamental difference between them.

through examples in each of these areas in the present paper. For some, I have done so elsewhere (e.g. Sawyer 2014); for others, I hope to do so in the future. The aim of the present paper is to provide an outline of the general framework in play.

## II

*Meaning Shift: Type 1 Cases.* I begin with a closer look at meaning shift of the first kind—meaning shift that is accompanied by a corresponding change in subject matter. Consider once again the term ‘meat’. Let us assume that at some earlier point in time,  $t_1$ , the term ‘meat’ meant *food in general* and that at the current point in time,  $t_2$ , the term means *animal flesh that is eaten as food*. Under this assumption, the meaning of the term ‘meat’ has clearly changed between then and now. Equally clearly, the extension of the term has changed.<sup>4</sup> There are some things, such as apples, bread and cake, which fall into the extension of the term at  $t_1$  but do not fall into its extension at  $t_2$ ; apples, bread and cake are food in general, but are not animal flesh that is eaten as food. This also marks a change in subject matter. An explanation is needed, then, of the change in meaning, extension and subject matter. It is provided by an account of linguistic meaning together with an account of the relation between meaning and extension, as follows.

I take meaning to be determined by use, where the use in question is located at the level of the linguistic community viewed as a whole rather than at the level of the individual speaker. Linguistic meaning in this sense is what dictionaries aim to record. I do not deny that there are individual aspects to meaning, but the phenomenon of meaning shift is best understood as a shift in conventional linguistic meaning in the communal sense, and it is, accordingly, conventional linguistic meaning that is the focus of the paper.

The way in which use determines linguistic meaning is complex. The linguistic meaning of a term cannot be understood as a simple aggregate of individual uses. Given what is often widespread diversity in use, a conjunction of individual uses

---

<sup>4</sup> Meaning shift does not necessitate a change in extension. However, many cases of meaning shift do involve a change in extension, and I will focus exclusively on these in the paper.

would tend towards an inconsistent meaning satisfied by nothing, and a disjunction of individual uses would eliminate the possibility of substantive error and undermine the normativity of language. Nor can the linguistic meaning of a term be understood in purely statistical terms, since a linguistic norm cannot be understood as a statistical norm given the merely descriptive nature of the latter. Each of these approaches ignores the fact that language is essentially a cooperative, social phenomenon and that linguistic meaning is shaped by patterns of deference amongst members of a linguistic community some of whom are more competent in the use of terms than others.

I suggest that the linguistic meaning of a term at a time be understood as the characterization of the relevant subject matter that members of the linguistic community would settle on at that time were they to reach reflective equilibrium in the context of a dialectic. For an account of linguistic meaning along these lines see Burge (1986) and (1989). See also Sawyer (2007). The account accommodates the obvious truth that some members of a linguistic community will be more competent in the use of a term than others, since in the context of a dialectic, those who are less competent would naturally defer to those who are more competent.<sup>5</sup> The account of linguistic meaning also accommodates the truth in Quine's claim (Quine 1951) that there is no separating truths of meaning from matters of fact, since the characterization of a subject matter could not be agreed without reflection on how the relevant term is applied in actual cases. The agreed characterization would reflect the actual practice of the linguistic community at the time, not by accommodating each individual's actual use, but by accommodating the actual use of the most competent and the deferential patterns amongst all. This is the sense in which linguistic meaning is determined by use. The agreed characterization would also set the relevant linguistic norms by establishing how the relevant term ought to be used. On this understanding, a term is used correctly if and only if it is used in accordance with the characterization of the subject matter that would be settled on in the context of a dialectic at the time.<sup>6</sup>

---

<sup>5</sup> See Putnam (1973) for a seminal discussion of the contribution of deference to linguistic meaning, although, for reasons given below, I disagree with the particular account of deference he provides. See also Burge (1982).

<sup>6</sup> Note that my claims about the normativity of language can be reconciled with the arguments against the normativity of language in Hattiangadi (2007), but space precludes discussion of the matter. I take

The notion of a dialectic here is an idealisation, but one that is intended to capture truths about the way a word is actually used at a given moment in time. This means that it must be understood as involving restrictions of at least two sorts. First, for terms which have an empirical element, such as ‘whale’, ‘atom’ and ‘sofa’, the admissible information must be restricted to include only empirical information which is accepted by the most competent speakers at the time. Second, for terms which have an apriori element, such as ‘number’, ‘validity’ and ‘knowledge’, the admissible level of reflection must be restricted to include only reflection available from within the confines of theories held by the most competent speakers at the time. In practice, terms will typically involve both elements. The reasons for the restrictions should be clear, since a violation of either restriction could lead to a settled characterization that not only outstrips current usage, but also tends towards an idealised characterization of a subject matter, thereby undermining the distinction between actual linguistic meaning and ideal linguistic meaning. The relevant notion, then, is not of maximal dialectical reflection on the subject matter, but of full reflection within actual empirical and theoretical boundaries.

This understanding of linguistic meaning provides an adequate explanation of the difference in meaning between the term ‘meat’ at  $t_1$  and the term ‘meat’ at  $t_2$ . The meaning of the term ‘meat’ at  $t_1$  is given by the characterization of the relevant subject matter that would have been settled on by members of the linguistic community at  $t_1$ ; the meaning of the term ‘meat’ at  $t_2$  is given by the characterization of the relevant subject matter that would be settled on by members of the current linguistic community, at  $t_2$ . These would clearly differ: the characterization of the subject matter associated with the term ‘meat’ at  $t_1$  would plausibly be *food in general*, while the characterization of the subject matter associated with the term ‘meat’ at  $t_2$  would plausibly be *animal flesh that is eaten as food*. The linguistic meaning of the term ‘meat’ at  $t_1$  differs from the linguistic meaning of the term at  $t_2$  because the underlying linguistic practices at those times differ.

---

thought, rather than language, to be the locus of fundamental, irreducible norms. A similar position is defended by Burge (1986).

The understanding of linguistic meaning I have suggested provides an adequate explanation of difference in meaning generally. Where there is change in linguistic practice, there will be change in linguistic meaning; and where there is change in linguistic meaning, this will be because there is change in linguistic practice. The meanings of terms such as ‘meat’, ‘clue’ and ‘spinster’, all of which display meaning shift of the first kind, have changed over time because the use of these terms in the linguistic community has changed over time. That is, the characterizations of the relevant subject matter that would be settled on by the relevant linguistic communities in the context of a dialectic would differ at the different times. The next task is to explain the change in extension of such terms over time.

The extension of a general, non-indexical term, such as ‘meat’, is typically understood as its range of applicability, where this in turn is understood as the class of entities to which the term correctly applies. However, given the distinction I draw later between the linguistic meaning of a term and the concept it expresses, this formulation is unacceptably ambiguous. Does the term correctly apply in virtue of its linguistic meaning, or in virtue of the concept it expresses? To avoid this ambiguity, I will say that the extension of a general, non-indexical term is the class of entities that satisfy the term’s linguistic meaning. The relation of satisfaction here is appropriate because the meaning of a general term is descriptive. Thus it is because an apple satisfies the description ‘food in general’ that it falls into the extension of the term ‘meat’ at  $t_1$ ; and it is because an apple does not satisfy the description ‘animal flesh that is eaten as food’ that it does not fall into the extension of the term ‘meat’ at  $t_2$ . The extension of a general term is specifically the extension of its linguistic meaning. For such terms, we can say that meaning determines extension.<sup>7</sup>

The account of the relation between meaning and extension I have suggested provides a general explanation of the change in extension that accompanies meaning

---

<sup>7</sup> I take the extension of a general term at time  $t$  to include all and only those entities that satisfy, at any point in time, the term’s meaning at  $t$ . For example, the extension of the term ‘meat’ at  $t_1$  includes all the apples, bread, cake and so on that have existed, do exist or will exist—not just those that exist at  $t_1$ . This means that, for such terms, the extension does not change over time unless the meaning of the term shifts. I leave to one side the question of whether it is determinate at a given time what will exist at a later time, and hence leave to one side related issues concerning determinacy of extension.



shift of the first kind. The extensions of terms such as ‘meat’, ‘clue’ and ‘spinster’ change over time because different entities satisfy their descriptive meanings at different times. Meaning shift of this first kind is also accompanied by a change in subject matter. What makes meaning shift of this first kind relatively straightforward is the fact that the intuitive change in subject matter of the relevant terms is marked, in each case, by the change in extension that results from the change in linguistic meaning. As such, the account of linguistic meaning and the relation between meaning and extension provides an adequate account of meaning shift of this first kind.<sup>8</sup>

### III

*Meaning Shift: Type 2 Cases.* Let us now take a closer look at meaning shift of the second kind—meaning shift that is not accompanied by a corresponding change in subject matter. Consider once again the term ‘whale’. Let us assume that at some earlier point in time,  $t_1$ , the term ‘whale’ meant *a very large fish with a streamlined hairless body, a horizontal tail fin, and a blowhole on top of the head for breathing* and that at the current point in time,  $t_2$ , the term means *a very large marine mammal with a streamlined hairless body, a horizontal tail fin, and a blowhole on top of the head for breathing*. Under this assumption, the term ‘whale’ has a different linguistic meaning at  $t_1$  from the linguistic meaning it has at  $t_2$ . The term ‘whale’ exhibits the phenomenon of meaning shift.

This fits naturally with the account of linguistic meaning given in the previous section. The characterizations of the relevant subject matter—whales—that would be settled on by the relevant linguistic communities in the context of a dialectic would differ at the different times. The account of linguistic meaning provided, then, explains the difference in meaning between the term ‘whale’ at  $t_1$  and the term ‘whale’ at  $t_2$ .

---

<sup>8</sup> Although the linguistic meaning of a term is specified descriptively, there will be an ineliminable demonstrative element to linguistic meaning because of the role that examples play in use and in the context of the dialectic.

But what explains the stability of the subject matter? It is natural to think that the subject matter of a term is, just like the extension of a term, determined by its linguistic meaning—that the subject matter of a term and the extension of that term go hand in hand (perhaps, even, that they are identical). These natural thoughts are consistent with cases that exhibit the first kind of meaning shift. But they are inconsistent with cases that exhibit the second kind of meaning shift. In such cases, the extension of a term changes along with a change in linguistic meaning despite the fact that the subject matter remains the same. That the extension changes can be seen by reflecting on the way in which linguistic meaning determines reference.

The extension of a general, non-indexical term, recall, is the class of entities that satisfy the term's linguistic meaning. The extension of the term 'whale' at  $t_1$ , then, is the class of entities that satisfy the description 'a very large fish with a streamlined hairless body, a horizontal tail fin, and a blowhole on top of the head for breathing'; and the extension of the term at  $t_2$  is the class of entities that satisfy the description 'a very large marine mammal with a streamlined hairless body, a horizontal tail fin, and a blowhole on top of the head for breathing'. On the assumption that the descriptions are not satisfied by the same class of entities—on the assumption, that is, that no fish is a mammal—it follows that the extension of the term 'whale' has changed.<sup>9</sup>

The extension of a term is determined by its descriptive linguistic meaning. Since a change in both the meaning and the extension of a term is consistent with no change in subject matter, we must conclude first, that the linguistic meaning of a term does not determine its subject matter, and second, that the subject matter of a term cannot be specified by a specification of the entities that fall into its extension, since the two may diverge.<sup>10</sup> We are in need, then, of an explanatory account of what

---

<sup>9</sup> The assumption may seem obvious, but see Dupré (1999) for an argument to the contrary. For this reason, the example may be ill-chosen. I have stuck with the example nonetheless because of its simplicity. My general framework is unaffected.

<sup>10</sup> The extent to which the extension of a term and the associated subject matter may diverge is considerable. In extreme cases—indeed, plausibly in the case of 'whale' as used at  $t_1$ —the extension of a term may be empty. This will no doubt seem counterintuitive. However, its counterintuitive appearance can be traced to a conflation between the linguistic meaning of a term and the concept expressed by that term—more specifically, a conflation between the extension of the linguistic meaning of a term and the extension of the referent of the concept expressed by the term. This is discussed in more detail below.

determines the intuitive notion of a subject matter given that it cannot be linguistic meaning. I will argue that the subject matter of a term is determined by the concept the term expresses.

To underscore the claim that linguistic meaning does not determine subject matter, note that identity of linguistic meaning is neither necessary nor sufficient for identity of subject matter. The failure of necessity is demonstrated by the example above of the term ‘whale’; the meaning of the term ‘whale’ at  $t_1$  is different from the meaning of the term ‘whale’ at  $t_2$ , and yet their subject matter—whales—is identical. The failure of sufficiency can be demonstrated by a counterfactual variant of the case. We need only imagine a counterfactual scenario in which the meaning of the term ‘whale’ at  $t_1$  remains fixed, because the practices of the linguistic community at  $t_1$  remain fixed, but the world is different in the following respect. In the counterfactual scenario the very large creatures with streamlined hairless bodies, horizontal tail fins, and blowholes on top of their heads for breathing are fish rather than mammals. The meaning of the term ‘whale’ at  $t_1$  is hypothesized to be the same in the counterfactual scenario as in the actual world, and yet the subject matter is clearly different; in the counterfactual scenario the subject matter is not whales (we may even assume that there are none)—it is fish of a certain kind. Identity of linguistic meaning, then, is neither necessary nor sufficient for identity of subject matter.

I have argued that the linguistic meaning of a term does not determine its subject matter on the grounds that identity of linguistic meaning is neither necessary nor sufficient for identity of subject matter. I have also claimed that it is the concept expressed by a term rather than its linguistic meaning that determines its subject matter. In the next section, I offer an account of concepts that supports the claim.

#### IV

*Concepts.* I take concepts to be mental representations that are components of thought, and I advocate an externalist understanding of mental representation according to which mental representation depends constitutively on non-representational relations between a subject and her environment. Concepts, then, are individuated, in part, but essentially, by relations between a thinker and objective

properties in her world. We can say that the property  $F$  that in part individuates a thinker's concept  $C$  is both the property to which  $C$  refers and the property that  $C$  represents.

This externalist understanding of concepts is illustrated by considering, once again, the example of the term 'whale'. On an externalist understanding of mental representation, the term 'whale' expresses the concept *whale* in the actual world both at  $t1$  and at  $t2$  in virtue of the fact that the linguistic community at  $t1$  and the linguistic community at  $t2$  both stand in the requisite relation to whales. In the counterfactual scenario, in contrast, the term 'whale' does not express the concept *whale* since, in the counterfactual scenario, the creatures that look and behave like whales are not whales but are instead a type of fish. We can call these creatures 'whale-fish'. In the counterfactual scenario, then, according to the externalist, the term 'whale' expresses the distinct concept *whale-fish*, since members of the linguistic community stand in the relevant relation not to whales but to whale-fish.

Note that the concept expressed by the term 'whale' in the actual world remains constant through the change in its linguistic meaning and hence through the change in its extension. As such, members of the linguistic community at  $t1$  and members of the linguistic community at  $t2$  are able to think and talk about whales despite the difference in linguistic meaning and extension of the term 'whale' at the different times. It is thought—specifically concepts—that provides stability through linguistic change. The externalist understanding of concepts, then, explains the fact that the subject matter of the term 'whale' is the same at  $t1$  as at  $t2$ . It is the concept expressed by a term that determines its subject matter. This is, in effect, a consequence of the fact that the subject matter itself enters into the individuation conditions of the relevant concept.

This marks a significant difference between language and thought that is often overlooked in the literature on externalism, although the distinction is clearly articulated by Burge (1986). In the linguistic case, we have a distinction between an individual's use of a term and the use of that term by the community taken as a whole. That is, we have a distinction between the characterization of a subject matter that an individual would provide, and the characterization of that subject matter that the

members of her linguistic community would settle on in the context of a dialectic. It is the latter that determines linguistic meaning. This explains how a term can have a single linguistic meaning at a time and yet nonetheless be used differently by different individuals at that time. The linguistic meaning of a term can be stable across different individual uses within a linguistic community because the linguistic meaning of a term is a function of how the community uses the term rather than a function of how the individual uses the term.

In the conceptual case, we have a corresponding distinction between an individual's conception, and the conception of the community viewed as a whole. That is, we have a distinction between the set of beliefs that an individual associates with a concept and the set of beliefs that the members of a community would agree were associated with the concept in the context of a dialectic. But in the conceptual case, neither the former nor the latter determine the concept expressed by a term. According to the externalist, concepts are not individuated by individual conceptions—they are not individuated by the way the individual thinker takes the world to be. Nor, according to the externalist, are concepts individuated by communal conceptions—they are not individuated by the way the community as a whole takes the world to be. Concepts are individuated by relations to objective properties independently of our individual or communal conceptions of them. Consequently, not only can a term express a single concept while being associated with different individual conceptions by different individuals, but, crucially, a term can express a single concept while being associated with different communal conceptions by different linguistic communities. This is what allows concepts, and not linguistic meaning, to determine a subject matter that remains stable through linguistic change. Linguistic meaning is a function of use, which is a function of communal conceptions; concepts, in contrast, stand over and above even communal conceptions. My account of concepts is closely aligned to a particular understanding of Frege's notion of sense (*'sinn'*), where this is understood as a cognitive notion and not a linguistic notion. See for example Frege (1892) & (1918). For convincing arguments against the (commonplace) identification of Fregean sense and linguistic meaning, see Burge (1979b), (1991) and (2012). In my framework, the linguistic meaning of a term determines its extension while the concept expressed by a term determines its

referent. There is an element of truth, then, in Frege's claim that sense determines reference. But reference is best understood as a cognitive relation, not a linguistic one.

My account of the linguistic meaning of a term as a community's best attempt to characterise the relevant subject matter invites an interpretation of Putnam's original Twin Earth thought experiment that differs from the standard externalist interpretation. See Putnam (1973). See also Kripke (1972), which invites the same reinterpretation. The set-up is familiar. Whereas the watery liquid on Earth is water ( $H_2O$ ), the watery liquid on Twin Earth is hypothesized to be twin water (XYZ). For simplicity, let us assume that the molecular structure of water was not known to anyone in 1700 on Earth, but was known generally in 2000, and that the molecular structure of twin water was not known to anyone in 1700 on Twin Earth, but was known generally in 2000. According to Putnam, the difference between Earth and Twin Earth implies: first, that the term 'water' on Earth differs in meaning and extension from the term 'water' on Twin Earth, both in 1700 and in 2000; and, second, that neither term undergoes a change in meaning or a corresponding change in extension through this period. The account of linguistic meaning I have provided entails the rejection of both of these claims.

Given that the difference between water and twin water is evident only at the level of molecular structure, and that in 1700 no-one knew about molecular structure, the best attempts of the linguistic community on Earth to characterize water in 1700 would have been identical to the best attempts of the linguistic community on Twin Earth to characterize twin water at that time. The linguistic meaning of the term 'water' on both Earth and Twin Earth would have been something like *clear, colourless, potable liquid that fills the river and lakes and sustains life*. Since the extension of a term is the class of entities that satisfy its descriptive meaning, the extension of the term 'water' would also have been the same on Earth as on Twin Earth in 1700. I assume for the sake of simplicity here that extensions are classes of entities across possible worlds.

By the time we reach 2000 the term 'water' on Earth does indeed differ in meaning and extension from the term 'water' on Twin Earth. I agree with Putnam on this point, albeit for different reasons. In 2000, the term 'water' on Earth differs in

linguistic meaning from the term ‘water’ on Twin Earth because each of the relevant communities had sufficient knowledge of the molecular structure of water and twin water respectively that their linguistic practices differ. The characterization of water that the linguistic community on Earth would settle on in the context of a dialectic includes a description of its molecular structure, H<sub>2</sub>O; and the characterization of twin water that the linguistic community on Twin Earth would settle on in the context of a dialectic includes a description of its molecular structure, XYZ. It is because the linguistic practices on Earth differ from those on Twin Earth that the term ‘water’ on Earth differs in linguistic meaning from the term ‘water’ on Twin Earth. And it is because the different linguistic meanings are satisfied by different liquids that the terms also differ in extension. What we have, then, is a change in linguistic meaning and a corresponding change in extension from 1700 to 2000 both on Earth and on Twin Earth. Putnam’s claim that there is no meaning shift and no change in extension should be rejected.

Putnam’s mistake was to think that the intuitions relating to Twin Earth were intuitions concerning the nature of language. They are, in part, intuitions about the nature of thought. It is not just that what Putnam says about language can also be said about thought. This, of course, is one of Burge’s contributions to the externalist theory of thought that my account of concepts depends on. See Burge (1979a). It is, rather, that Putnam argues for claims about the nature of language on the basis of intuitions that are best understood as intuitions about the nature of thought. We can see this by noting a tension in Putnam’s account. Putnam distinguishes two elements that he claims are missing from traditional semantic theory. The first is ‘the contribution of society’; the second is ‘the contribution of the real world’ (Putnam 1973, p. 711). But the two elements pull in different directions.

The contribution of society to linguistic meaning implies that the meaning of the term ‘water’ as used on Earth at *t* is identical to the meaning of the term ‘water’ as used on Twin Earth at *t*. As we saw above, this follows from my account of linguistic meaning which emphasises the role of society in the determination of meaning. But the implication does not depend on my account. Putnam himself introduces the notion of linguistic deference, which he articulates as deference to ‘experts’. In this context, experts have typically been understood as those who are

knowledgeable about the subject matter. But the experts to whom we defer at a given moment in time are actual experts rather than ideal, all-knowing experts. They are not infallible and cannot be presumed to know everything there is to know about the relevant subject matter. The history of science is replete with experts who were wrong about at least some aspects of the subject matter that fell into their area of expertise. Experts, then, insofar as they are actual members of a linguistic community to whom people might defer, are merely those who are the most competent amongst us in the use of a term. If the linguistic meaning of an individual's term is determined by what the experts mean by that term, then linguistic meaning is determined by the experts' best attempt to characterize the subject matter.

The contribution of the real world to linguistic meaning, however, implies, as Putnam maintains, that the meaning of the term 'water' as used on Earth at *t<sub>1</sub>* *is different from* the meaning of the term 'water' as used on Twin Earth at *t<sub>1</sub>*. Water and twin water are objectively different, independently of what anyone, even an expert, believes, and this apparently supports the intuition of difference in linguistic meaning. If the two contributions that Putnam identifies are understood as contributions to a single phenomenon—whether that be linguistic meaning, as Putnam supposed, or thought—they will generate inconsistencies. Since the experts in society are themselves prone to error in their characterizations of the real world, the contribution of society and the contribution of the real world will sometimes deliver inconsistent results about the nature of the single phenomenon in question; society will pull towards change while the real world pulls towards stability.

To resolve the problem, the two contributions need to be seen as contributions to different phenomena. The contribution of society that Putnam identifies is best understood as a contribution to language, and the contribution of the real world that he identifies is best understood as a contribution to thought. This is what I have been suggesting throughout the paper. The linguistic practices of society constrain linguistic meaning, which determines extension. The real world constrains concepts, which determine reference. On this understanding, we can make sense of Putnam's (incorrect) claim that the term 'water' on Earth does not change its meaning over time by reconstruing it as the claim that the concept expressed by the term does not change over time. We can then do justice to both of Putnam's claims: although the term



‘water’ has the same linguistic meaning in 1700 on Earth as it does on Twin Earth, the subject matter differs; moreover, although the term ‘water’ on Earth in 2000 differs in both linguistic meaning and extension from the term ‘water’ on Earth in 1700, the subject matter remains the same. The term ‘water’ exhibits meaning shift of the second kind.

Linguistic meaning is determined by patterns of actual use; concepts are determined by real relations to objective properties. As such, concepts can provide stability through linguistic change. This provides an adequate explanation of meaning shift of the second kind.

## V

*Error, Disagreement and Truth.* Distinguishing clearly between linguistic meaning and concepts provides us with a theoretical framework in which we can make sense both of change and of stability. Linguistic meaning tracks the former; concepts explain the latter. Without an account of stability, all meaning shift is meaning shift of the first kind and hence all meaning shift will inevitably amount to changing the topic. This brings with it the kind of incommensurability that Kuhn (1962) thought plagued theory change, and entails an unacceptable relativism.

Take the sentence ‘Bread is a form of meat’. The sentence was generally regarded as true at  $t_1$  but is generally regarded as false at  $t_2$ . Nonetheless, it is clear that any appearance of disagreement is merely superficial. The sentences simply express different thoughts at the different times, because the concept expressed by the term ‘meat’ at  $t_1$  is distinct from the concept expressed by the term ‘meat’ at  $t_2$ . This means that the sentence is true at  $t_1$  (true relative to the linguistic framework at  $t_1$ ) but false at  $t_2$  (false relative to the linguistic framework at  $t_2$ ). The notion of a linguistic framework is articulated in Carnap (1947), but see especially Carnap (1950). In cases of this first kind, there is no stability of subject matter because a change in linguistic meaning is tracked by a change in concept.

Now consider the sentence ‘Whales are fish’. This sentence was also generally regarded as true at  $t_1$  but is generally regarded as false at  $t_2$ . If a change in linguistic

meaning is always tracked by a change in concept, then we are forced to say that here too, the sentences simply express different thoughts at the different times, because the concept expressed by the term ‘whale’ at  $t_1$  is distinct from the concept expressed by the term ‘whale’ at  $t_2$ , and, consequently, the sentence is true at  $t_1$  (true relative to the linguistic framework at  $t_1$ ) but false at  $t_2$  (false relative to the linguistic framework at  $t_2$ ). But once we distinguish between the linguistic meaning of a term and the concept expressed by a term in the way I have suggested, we can allow that the concept expressed by the term ‘whale’ remains constant through the linguistic change between  $t_1$  and  $t_2$ , and hence understand the apparent disagreement as a genuine disagreement about whales. The distinction between language and thought brings with it a distinction between linguistic frameworks, individuated by actual linguistic practices, and conceptual frameworks, individuated by real relations to the world. This provides an antidote both to Kuhn’s claims about incommensurability and to Carnap’s (1947) eschewal of so-called ‘external’ questions.

It also addresses head-on the apparent dilemma raised in Sainsbury’s (2014). According to Sainsbury, we are faced with a dilemma. Either the sentence ‘Whales are fish’ means the same in ancient times as now, or it doesn’t. To say that it does, according to Sainsbury, belies the fact that meaning is determined by use; but to say that it doesn’t belies the fact that there is substantive disagreement across the two times. Sainsbury concludes, ‘There is something paradoxical about fish.’ And conjectures that ‘[a]n adequate resolution will require carefully formulated metasemantic principles.’ (Sainsbury 2014, p. 5). The alternative I have proposed resolves the apparent dilemma. Once we distinguish the linguistic from the conceptual, we can assert both that the meaning of the sentence differs at the two times, which accommodates the fact that the use of the term differs at the two times; and that the sentence expresses the same thought at the two times, which accommodates the fact that there is substantive disagreement.

Reflection on the change in linguistic meaning of the term ‘whale’ over time also brings to light a point about the rationale for theoretical and linguistic change. I have said that the linguistic meaning of a term can be thought of as a community’s best attempt to characterise the relevant subject matter. We have seen that there is a potential gap between a community’s *best attempt* to characterize the relevant subject

matter and the *correct* characterization of that subject matter. And it is recognition of this gap in particular cases that drives theory change and hence change in linguistic meaning. The gap is identified in Burge (1986), where the difference between language and thought is also articulated. In the actual world, the linguistic community's best attempt to characterise whales at  $t_1$  turned out to be incorrect. Nonetheless, it provided the linguistic meaning of the term 'whale' at that time. The shift in meaning of the term 'whale' between  $t_1$  and  $t_2$  can, then, be understood as a change in linguistic practice brought about by a realization that the initial characterization was incorrect. In the context of a dialectic at some point between  $t_1$  and  $t_2$ , the initial characterization would have been rejected and replaced by one that was considered to be accurate. The characterization may, of course, still have fallen short, but this allows for further change in meaning as the community's knowledge evolves.

Our characterizations of various subject matters often fall short of the facts. And a change in linguistic practice can often be understood as an attempt to characterize the facts more accurately than has previously been achieved. These attempts rationalise meaning shift of the second kind. As such, wherever there is the possibility of less than full understanding of a given subject matter, there is the possibility of this second kind of meaning shift. Examples abound not just in empirical science, but in philosophy, mathematics, and logic, as well as in the social, political and moral realm. As I said at the outset, I do not have the space to discuss detailed examples in each case. For now, I offer some brief remarks on an example drawn from the recent literature on ameliorative change and conceptual engineering. The literature is vast, and growing, but see for example Appiah (1992), Haslanger (2012), Ecklund (2014) and Cappelen (2018).

I take as my example the term 'rape'. The debate about whether rape should be recognised as a phenomenon that could occur within marriage as well as outside of marriage goes back a long way. But the debate, I suggest, is correctly understood as a debate about rape. This can only be true if there is some stability of subject matter across the two sides of that debate. Linguistic practice surrounding the term 'rape' has clearly changed over time. This means that the linguistic meaning of the term 'rape' has changed over time. But we should not, I suggest, accept the kind of relativism

about rape that would follow from thinking of the change in meaning as meaning shift of the first kind. Acts of rape within marriage may not previously have been recognised as such, but they were acts of rape nonetheless. The change in linguistic practice reflects a recognition that the earlier linguistic practice got the facts wrong. It reflected a misunderstanding of the nature of rape. Naturally, there is more to be said both about the specific example and, more generally, about when to classify an apparent change in linguistic practice as exemplifying a change in linguistic meaning of the second kind as opposed to a shift in focus to a different, perhaps related aspect of the world that comes to seem more pressing. I will return to these issues on a future occasion.

For now, I finish with a potential objection to the overall framework. It might be thought that it is my account of linguistic meaning that should be rejected. Perhaps instead, we should understand linguistic meaning not as tied to actual practice, but as determined by real relations to objective properties in the way that I have understood concepts to be determined. On the proposed alternative view, the linguistic meaning of a term would determine its subject matter as well as its extension. Thus the subject matter and extension of the term ‘rape’ in the past would be identical to the subject matter and extension of the term now. However, this has the benefit of accommodating stability at the cost of eliminating meaning shift of the second kind. On this view, the meaning of the term ‘rape’ would be the same across time, and hence there would be no meaning shift. Moreover, it eliminates a theoretical notion of linguistic meaning that tracks the actual use of a term by a linguistic community. There are two important aspects to our language-use that need to be explained: actual linguistic usage and, as we might put it, ideal linguistic usage—how certain of our terms are in fact used and how those terms *ought* to be used. On this proposed alternative view of linguistic meaning, we fail to capture the first; but without a theory of concepts, we fail to capture the second. This distinction is discussed in Burge (2012, p. 589). Perhaps particularly in the context of moral terms, we should be aiming to use our terms in the way in which they ought to be used, not just in the sense of according with the community’s best attempt to characterize the subject matter, but in the sense of according with the *correct* characterization of the subject matter. I take ‘rape’, for example, to be a moral term because it is essentially a term

for an act that ought, morally, not to occur. Concepts provide us with the stability of subject matter that makes the move from the one to the other possible.

## VI

*Conclusion.* I have argued that the linguistic meaning of a term is not to be equated with the concept expressed by that term because the two have different explanatory roles and are subject to different norms. Where there is stability of subject matter through linguistic change, it is concepts that explain the stability. It is also concepts that explain the possibility of genuine disagreement over a single subject matter, thus avoiding widespread linguistic and conceptual relativism. Concepts provide us with a fundamental representational relation to aspects of reality even when we do not fully comprehend them. Our lack of comprehension is evident in almost every area of inquiry: philosophy, mathematics, logic, empirical science, the social, the political and the moral. But in all these areas, we can strive for understanding and for meaning shift in the right direction. We can strive, that is, towards the true and the good.<sup>11</sup>

*Department of Philosophy*  
*University of Sussex*  
*Brighton BN1 9QN*  
*s.a.sawyer@sussex.ac.uk*

## REFERENCES

- Appiah, Kwame Anthony. 1992: *In My Father's House: Africa in the Philosophy of Culture*. Oxford: Oxford University Press.
- Bromberger, Sylvain. 2011: 'What are Words? Comments on Kaplan (1990), on Hawthorne and Lepore, and on the issue'. *Journal of Philosophy*, 108, pp. 486-503.

---

<sup>11</sup> Thanks to John Dupré and Guy Longworth for comments and discussion. Thanks also for insightful questions from those who attended the Aristotelian Society meeting at which this paper was originally delivered.

- Burge, Tyler. 1979a: 'Individualism and the Mental'. *Midwest Studies in Philosophy*, 4, pp. 73-121.
- 1979b: 'Sinning Against Frege'. *Philosophical Review*, 88 (3), pp. 398-432.
- 1982: 'Other Bodies'. In A. Woodfield (ed.), *Thought and Object: Essays on Intentionality*. Oxford: Clarendon Press, pp. 97-120.
- 1986: 'Intellectual Norms and Foundations of Mind'. *Journal of Philosophy*, 83 (12), pp. 697-720.
- 1989: 'Wherein is Language Social?' In Alexander George (ed.), *Reflections on Chomsky*. Oxford: Blackwell, pp. 175-92.
- 1991: 'Frege on Sense and Linguistic Meaning'. In D. Bell & N. Cooper (eds), *The Analytic Tradition: Meaning, Thought and Knowledge*. Oxford: Blackwell, pp. 30-60.
- 2012: 'Living Wages of Sinn'. *Journal of Philosophy*, 109, pp. 40-84.
- Cappelen, Herman. 2018: *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.
- Carnap, Rudolf. 1947: *Meaning and Necessity. Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press.
- 1950: 'Empiricism, Semantics and Ontology'. *Revue Internationale de Philosophie*, 4, pp. 20-40.
- Dupré, John. 1999: 'Are Whales Fish?' In D. L. Medin & S. Atran (eds), *Folkbiology*. Cambridge, Mass.: MIT Press, pp. 461-76.
- Eklund, Matti. 2014: 'Replacing Truth'. In A. Burgess & B. Sherman (eds), *Metasemantics: New Essays on the Foundations of Meaning*. Oxford: Oxford University Press, pp. 293-310.
- Frege, Gottlob. 1892: 'On Sense and Reference'. Translated 1948 in *Philosophical Review*, 57 (3), pp. 209-30.
- 1918: 'The Thought: a Logical Inquiry'. Translated 1956 in *Mind*, 65 (1): 289-311.
- Haslanger, Sally. 2012: *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.
- Hattiangadi, A. 2007: *Oughts and Thoughts: Rule Following and the Normativity of Content*. Oxford: Oxford University Press.
- Hawthorne, John. & Lepore, Ernest. 2011: 'On Words'. *Journal of Philosophy*, 108, pp. 447-85.

- Kaplan, David. 1990: 'Words'. *Aristotelian Society Supplementary Volume*, 64, pp. 93-119.
- 2011: 'Words on Words'. *Journal of Philosophy*, 108, pp. 504-29.
- Kripke, Saul. 1972: *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
- Kuhn, Thomas. 1962: *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.
- Putnam, Hilary. 1973: 'Meaning and Reference'. *Journal of Philosophy*, 70, pp. 699-711.
- Quine, W.V.O. 1951: 'Two Dogmas of Empiricism'. *Philosophical Review*, 60, pp. 20-43.
- Sainsbury, Mark. 2014: 'Fishy Business'. *Analysis*, 74, pp. 3-5.
- Sawyer, Sarah. 2007: 'There is No Viable Notion of Narrow Content'. In B. McLaughlin & J. Cohen (eds), *Contemporary Debates in the Philosophy of Mind*. Oxford: Blackwell.
- 2014: 'Minds and Morals'. *Philosophical Issues, Extended Knowledge*, 24, pp. 393-408.