Referentialism and Internalism

Antony Eagle

antonyeagle.org

Philosophy of Language » Lecture 2

Contents

- Referentialism
- Referentialism and Propositions
- Internalism
- The Nature of Concepts

Where we are

- It's almost impossible to come up with definitions in the strict sense – **analytically true necessary and sufficient conditions** – for ordinary non-technical terms.
  - The two parts pull against each other – the more accurate the proposed necessary and sufficient conditions, the less plausible that all competent speakers implicitly know them.
- Dictionary definitions are looser, being instructions to help speakers who are already mostly competent join in our existing **convention** for **using** an expression.
-
  - Understanding a word might be better thought of as a matter of being disposed to use it in accordance with community standards. That is why dictionary definitions are useful: linguistic competence is about use.
- Definitions, being linguistic entities, are poor candidates to be meanings, since the word-world relationship is left unconstrained, no matter how many word-word relations we specify.
- But that still leaves open a central question: **what are meanings?**

Referentialism

The Simplest Story: Referentialism

- The simplest story about the word-world connection is the most direct.
- Everyone agrees that (some? all?) words **refer**: that the word *grass* refers to grass, and the word *Antony* refers to me.
- And everyone agrees that the meaning must **fix** the referent. It is something about what *Antony* means that enables it to refer to me.
- So we can frame our question about the word-world connection: **what is a meaning, such that the meaning of a word determines its referent**?
- The **referentialist** answer is simple: the meaning (intension) of a word **just is** its referent.

> the word *Iceland*, for example, has as its meaning that very island, a huge chunk of rock and ice in the northern Atlantic Ocean. … the meaning of the word just is what it picks out in the world. (Elbourne 2011: 14–15).

Referentialist Theories of Understanding

- The referentialist thinks that the meaning of a word is its referent, so something **not** in the head. Coherent use, as long as it manages to have an extension which is close to that of a real object or property, is sufficient to generate meaning.
- In the last lecture I suggested, following Wittgenstein (1967) and Williamson (2006), that to understand an expression is to **use it correctly**, i.e., in circumstances appropriate to its actual referent, as we've already sketched.
- That ability to successfully participate in community use is almost certainly grounded in something in the head – concepts, or dispositions, or mental representations of some sort.
- But, according to the referentialist, **those** aren't meanings – while the deployment of concepts is how we implement our meaningful use of language, they aren't what language is about.
  - Analogy: our retinal state correlates with the external objects in our visual field. But it is nonsense to say that **what we see** is our own retina – the object of sight is the external tree, and the object of the expression *that tree* is likewise the external tree.

Millianism: Referentialism about proper names

> Proper names are not connotative: they denote the individuals who are called by them; but they do not indicate or imply any attributes as belonging to those individuals. When we name a child by the name Paul, or a dog by the name Cæsar, these names are simply marks used to enable those individuals to be made subjects of discourse. It may be said, indeed, that we must have had some reason for giving them those names rather than any others; and this is true; but the name, once given, is independent of the reason. A man may have been named John, because that was the name of his father; a town may have been named Dartmouth, because it is situated at the mouth of the Dart. But it is no part of the signification of the word John, that the father of the person so called bore the same name; nor even of the word Dartmouth, to be situated at the mouth of the Dart. If sand should choke up the mouth of the river … the name of the town would not necessarily be changed.… Proper names are attached to the objects themselves, and are not dependent on the continuance of any attribute of the object. (Mill 1882: bk. 1, chapter 2, §5)

Referents For All Expressions?

- Referentialism is intuitively plausible enough for proper names, and for kind terms (*gold*, *grass*, etc.).
- But what could be a referent for a **predicate**: for the *is green* part of *grass is green*?
- The standard answer: a predicate denotes a **property**. The predicate *is green* denotes the property greenness; the relational predicate *is taller than* denotes the taller-than relation.
- So the meaning of *chair* is not some collection of necessary and sufficient conditions; it is the **property of chairhood**, which of course is had by all and only chairs, whether or not we **know** which specific things those are, or can give an **analysis** of that property.

Extensions, Intensions, and Properties

- What are properties? Elbourne (2011: 16–18) seems to be a bit worried about the metaphysics of properties: the traditional question of **universals** (Armstrong 1989). (He seems particularly exercised by their being multilocated wherever they are instantiated.)
- But arguably there must be **something** that plays the **property role** – grounding similarity between things, differentiating gerrymandered from natural classifications, etc. – and whatever that is, it can be an external referent for a predicate.
- One very minimal theory of properties appeals to the notion of an intension from last lecture, and proposes to identify a property with a certain kind of set:

> The property of being a donkey, for instance, is the class of *all* the donkeys. This property belongs to – this class contains – not only the actual donkeys of this world we live in, but also all the unactualised, otherworldly donkeys. (Lewis 1983: 344)

- I don't propose to litigate these disputes over the nature of properties; our working hypothesis will be that **any** adequate theory will determine an intension and that is all the referentialist ultimately needs.

How Do Words Get Referents?

- If the meanings of words are just 'out there', things and classes of things, how do words and their meanings get matched up?
- One standard referentialist answer (not the only answer):
  1. Our **collective use** of the expression determines – at least when cleaned up to account for performance errors – an extension, the things we've applied the expression to;
  2. The referent is then the **best candidate** object/property in the vicinity of that extension.
- We collectively have a certain pattern of uses of a word. Some of those uses are **mistakes**, which we don't count. For example, applications of a term in poor conditions – bad light, while drunk or otherwise impaired, etc.
- There might be **borderline cases** (think of colour terms), where some speakers might apply the term and others withhold it, but no speaker thinks it is a mere mistake to do either. The **determinate** extension of a term will be those things to which all competent non-mistaken speakers apply the term.
- Then the referent of the word is the best or closest candidates to that determinate extension.

Best Candidate Referents

- What is 'best' here (Speaks 2021: §3.2)?
  1. Maybe it is the referent which is **causally responsible** for our extension-generating use; or
  2. Maybe it is the **most natural** or **eligible** object/property in the vicinity of that extension: 'Reference consists in part of what we do in language or thought when we refer, but in part it consists in eligibility of the referent' (Lewis 1983: 371).
- These different accounts suit different expressions, as we will see.
  - So proper names of ordinary material things might well refer to the thing which ultimately causes the use of that name, as the **causal theory of reference** suggests – a topic to which we will return.
  - But properties, and some associated common nouns, might be better understood as having their referent fixed by eligibility. The word *chair* denotes the property of being a chair, which is the most natural grouping in the vicinity of the determinate extension established by our use. Terms with borderline cases – maybe even *chair* itself! – might be understood as having a number of equally eligible candidate meanings capturing our use.

Meaning and Co-reference

- We've been assuming so far that the referentialist takes the meanings of words to be physical things, or classes of them.
- But as Frege notes, names for the same thing can appear to differ in meaning, making **different contributions** to sentences in which they appear.
- E.g., these two appear to differ in meaning, despite (2) resulting from (1) by substitution of a **co-referring** term:
  (1) Superman is Superman
  (2) Superman is Clark Kent
- Frege concludes:

> with a proper name, it depends on how whatever it refers to is presented. … The different thoughts which thus result from the same sentence correspond in their truth-value, of course…. Nevertheless their distinctness must be recognized. So it must really be demanded that a single way in which whatever is referred to is presented be associated with every proper name. (Frege 1918: 298)

Sense and Referentialism

- Frege uses 'referent' to mean the thing a name denotes, but here is arguing that another dimension of meaning should also be associated with names, the **mode of presentation** or **sense**.
  - Frege's distinction, and his grounds for it, are an important topic in our fourth lecture.
- But what matters today is that Frege's theory of senses is a referentialist one, because he takes these expressions to pick out non-linguistic entities that belong to what he calls a 'third realm' (Frege 1918: 302).
  - These entities are like ideas in that they 'cannot be perceived by the senses', but unlike ideas they exist independently of any consciousness or awareness, 'like a planet which, already before anyone has seen it, has been in interaction with other planets' (Frege 1918: 302).
  - There is a query about whether Frege should be counted a referentialist – after all, he denies the Millian thesis that the meaning of a name is its referent! – but as a non-internalist with idiosyncratic views about the metaphysics of meaning, he might bemo treated as an atypical referentialist.
- And Frege's account is especially well-suited to integrate with his referentialist account of the meanings of sentences – to which we turn next.

Referentialism and Propositions

Propositions and Thoughts

- A thoroughgoing referentialism offers meanings for all expressions. So what about referents for **sentences**?
- The basic idea is that the meaning of a sentence is a **proposition**, the kind of entity Frege (1918) calls a 'thought':

> I call a thought something for which the question of truth arises. So I ascribe what is false to a thought just as much as what is true. … The thought, in itself immaterial, clothes itself in the material garment of a sentence and thereby becomes comprehensible to us. We say a sentence expresses a thought. (Frege 1918: 292)
>
> I … recognize the thought, which other people can grasp just as much as I, as being independent of me. (Frege 1918: 307)

- Frege's terminology is unfortunate – he strenuously distances himself from the suggestion that thoughts are ideas in the mind (Frege 1918: 302, 308), but if so the term *thought* is perhaps misleading. I will thus prefer *proposition*.
  - Though some uses of *thought* fit Frege's pattern, e.g., *Are you thinking what I'm thinking?*

The Proposition Role

- Frege's remarks emphasise two features of propositions:
  1. They are the primary **bearers** of truth and falsity; and
  2. They are the **meanings** of sentences – principally **declarative** sentences which state how things are (Frege 1918: 293).
- To these is usually added a third and/or fourth, not clearly distinct:
  3. They are the **objects** of our mental attitudes;
  4. They are the entities grasped or apprehended in thinking (Frege 1918: 294).
- A proposition is typically denoted by a *that*-clause; in this example, the underlined clause refers to a proposition.

  (3) Antony believes <u>that Sylvester is at the park</u>.
- The verb *believes* – along with other verbs taking a *that*-clause complement like *knows*, *asserts*, etc. – expresses a **propositional attitude**.
- Propositions are the main focus of lecture 5.

Russellian Propositions

- The simplest idea, often ascribed to Russell (1903), is that propositions are structured entities, constructed out of the meanings of the constituent words.
- So the meaning of *grass is green*, the proposition it expresses, might be characterised by the structured entity $\langle \text{greenness}, \text{grass} \rangle$, if the referents of words are things and properties, or perhaps by the senses of those things if Frege is right.
  - The angle brackets signify an **ordered pair**: this is a mathematical entity with various properties, the only interesting one for our purposes being that the order matters, so that $\langle x, y \rangle$ is generally different from $\langle y, x \rangle$.
  - Order matters: for Russell, the propositions expressed by *Jane loves Joan* and *Joan loves Jane* consist of the same entities, so much be differentiated by how they are put together: e.g., $\langle \langle L, x \rangle, y \rangle$ vs. $\langle \langle L, y \rangle, x \rangle$.
- Even if the entities involved are physical, an ordered pair is an abstract mathematical entity, so regardless of whether Frege is right that word meanings inhabit the 'third realm', sentence meanings appear to.
- This is partly because there are false propositions, and these ought **not** correspond to any realized or actual **fact**.

Empty Names and Abstract Objects

- A problem for simple referentialism: what is the meaning of so-called **empty names** like *Santa Claus* and *Pegasus* (Elbourne 2011: 19)?
- They have no physical referent – so how can they have meaning, or make any contribution to the proposition expressed by sentences in which they appear?
  - Moreover, it seems that all empty names will have the same lack of meaning, making them all synonymous.
- Yet there must be some proposition expressed by *Pegasus flies*, because someone could **assert** or **believe** it: mistakenly, perhaps, but sincerely.
- The usual proposal is to supply a referent for *Santa Claus*: an **abstract object**, maybe a fictional character (Elbourne 2011: 20–21), maybe a Fregean sense.
- But then we face a dilemma, for **negative existentials**, like *Santa Claus does not exist*, which should come out true, appear to come out false if we do supply an entity (of whatever sort) to be the meaning of the name (Elbourne 2011: 22).
- I don't think empty names should drive our semantic theorising; they are too exotic to be core cases. (Still less empty predicates.) But if there is another theory which is also plausible and which handles apparently referentless expressions better, perhaps that theory is on balance to be preferred.

Internalism

Internalism about Meanings

> On the other side of the floor, we have the advocates of the *internalist theory of meaning*. They suggest that word meanings are most fruitfully thought of as ideas or concepts in our heads. Take a concept, such as the concept I have of Iceland. It is some psychological entity. Ultimately, if we are correct to suppose that we do our thinking with our brains, this concept of mine is presumably a structure composed of cells inside my head. … Since the island of Iceland resembles or *falls under* this concept of mine, I use this concept to think about Iceland. And since the concept also forms part of a word (i.e., since it is the meaning of a word), I use that word, *Iceland*, to talk about Iceland. Inside your head, you presumably have a very similar word, so that when you hear me say 'Iceland' your concept of Iceland is activated. (Elbourne 2011: 15)

Internalism: Meanings are Concepts

- The **internalist theory of meaning** states that the meaning of a word is something internal to us – generally an idea or **concept**.
- So the meaning of *chair* (for me) is a mental entity, the concept CHAIR in my mind.
  - This is something like the view associated with Locke: that words stand for 'ideas' in the mind of the speaker.
- There is something compelling about this: after all, isn't it just phenomenologically obvious that when I speak, I give voice to my ideas? And so there must be some significant component of the meaning of my words which is fixed by, or grounded in, what's going on inside my head.
- Internalism tends to appeal to those who see linguistics as a branch of **cognitive science**: for it makes meanings available in the explanation of cognitive processes, e.g., reasoning.
  - The referentialist view agrees that mental entities like concepts are involved in reasoning; but meaning can only be a passenger in the explanation of how it works.

Concepts and Competence: Arguing for Internalism

- Reliable use of the word *chair* requires the concept CHAIR.
  - Their reliable use shows that they successfully use the concept to classify objects they encounter.
- When we acquire **competence** in the use of a term from a dictionary, or come to **understand** it through observing other people's use, those **abilities** are grounded in how I am and what I can do – they are features of my mental skill-set, not something external to me.
- Since (i) concept possession explains my use of an expression, and (ii) meaning is determined by use, it is natural to identify the meaning of an expression I use with the concept which grounds my correct use.
- The word-world relation is then **derivative** from the concept-world relation – linguistic representation is directly parasitic upon mental representation.

Sentence Meaning

- The internalist theory will say that the meanings of sentences are, just like the meanings of words, mental entities, which contain concepts as their constituents.
- They might even say that the meaning of a sentence is also a concept: not one that classifies ordinary objects into those falling under the concept and those that do not, but rather classifies **possible situations**, into those where the sentence is true and those where it is not.
  - A Fregean proposition is not a concept, but Frege may well agree that we **apprehend** a proposition just when we have this sort of concept that appropriately tracks the proposition – i.e., one we are prone to apply to just those situations in which the proposition is true (Frege 1918: 310).
- Note the **simplicity** of the internalist picture here: all meanings are concepts, simple or complex.
  - Compare the referentialist, who offers physical objects, properties, and abstract entities like structured propositions, so that meaning is not a unified category.

Chomsky's Internalism About Words

- An influential internalist view is associated with Chomsky (2000):

> Words … are mental entities that consist of three parts: phonological information…; syntactic information; and semantic information, or meaning…. As for the precise form that this semantic information takes, Chomsky has little to say; it is deeply mysterious. (Elbourne 2011: 23)

- So the meaning of a word is literally **part** of a word; and a word is a mental entity. Since it doesn't seem very plausible that you and I share our mental constituents, there isn't just one word *chair*: rather, there is a distinct word $chair_\alpha$ in the mind of each competent speaker α.
  - These words are very similar, functionally and qualitatively; but they are numerically distinct, and may even differ in their syntax and semantics.
- To be clear, this is an example: not every internalist theory of meaning follows Chomsky, though all agree with the part that says word meanings are mental entities.

The Plurality of Concepts

- By identifying meanings with mental entities, the internalist strictly denies that there is a unity of meanings between different speakers.
  - My concept CHAIR is not likely to coincide perfectly with your concept CHAIR.
  - The differences may be small, so that we won't be confused if we both use the expression *chair* to denote our individual concepts, but there is no such thing, strictly speaking, as **the** extension of *chair* in English – rather, there are the extensions of $\text{CHAIR}_{\text{Antony}}$, $\text{CHAIR}_{\text{Atheer}}$, etc.
- This gives rise to one of the more striking internalist theses: that external public language doesn't really exist, only internal conceptual structures or 'grammars':

> grammars have to have a real existence, that is, there is something in your brain that corresponds to the grammar. … But there is nothing in the real world corresponding to language. In fact it could very well turn out that there is no intelligible notion of language. (Chomsky 1982: 107)

The Privacy of Mental Representations

> My companion and I are convinced that we both see the same field; but each of us has a particular sense-impression of green. I notice a strawberry among the green strawberry leaves. My companion does not notice it, he is colour-blind. The colour-impression, which he receives from the strawberry, is not noticeably different from the one he receives from the leaf. Now does my companion see the green leaf as red, or does he see the red berry as green, or does he see both as of one colour with which I am not acquainted at all? These are unanswerable, indeed really non-sensical, questions. For when the word "red" does not state a property of things but is supposed to characterize sense-impressions belonging to my consciousness, it is only applicable within the sphere of my consciousness. … It is so much of the essence of each of my ideas to be the content of my consciousness, that every idea of another person is, just as such, distinct from mine. (Frege 1918: 299–300)

Content is Public and Interpersonal

> If the thought I express in the Pythagorean theorem can be recognized by others just as much as by me then it does not belong to the content of my consciousness, I am not its bearer; yet I can, nevertheless, recognize it to be true. However, if it is not the same thought at all which is taken to be the content of the Pythagorean theorem by me and by another person, … truth would be restricted to the content of my consciousness and it would remain doubtful whether anything at all comparable occurred in the consciousness of others.
>
> If every thought requires an owner and belongs to the contents of his consciousness, then the thought has this owner alone; and there is no science common to many on which many could work, but perhaps I have my science, a totality of thoughts whose owner I am, and another person has his. Each of us is concerned with contents of his own consciousness. No contradiction between the two sciences would then be possible, and it would really be idle to dispute about truth; as idle, indeed almost as ludicrous, as for two people to dispute whether a hundred-mark note were genuine, where each meant the one he himself had in his pocket and understood the word 'genuine' in his own particular sense. If someone takes thoughts to be ideas, what he then accepts as true is, on his own view, the content of his consciousness…. If he heard from me the opinion that a thought is not an idea he could not dispute it, for, indeed, it would not now concern him. (Frege 1918: 301–2)

Internalists Respond: Is Language Interpersonal?

- The requirement for public meanings seems to enforce a conformity that is not manifest in our use, as Elbourne (2011: 31) notes.
- There seems to be a need for private meaning that diverges from that of others: consider the phrase *lucked out*, which is synonymous with *lucky* in North American Englishes but which for many speakers of Australian English means *out of luck*.
  - Examples can be multiplied: *enormity* (a big thing vs. a grave sin), *begs the question* (raises the question vs. argues in a circle), *disinterested* (uninterested vs. neutral),….
- Of course the significance of this observation is disputed.
- Referentialists think: in these cases we have dispositions to use a term that vary across a population, and each of these stable dispositions fixes an external referent. The thesis that meaning is fixed by use, common to referentialists, faces no obstacle from these examples.
- On the other hand, internalists face a problem with deviant speakers whose usage isn't systematic or widespread. The first person to use *disinterested* to mean *uninterested* was just wrong, and misunderstood the contribution of *interest* to this expression – can the internalist explain that?

The Irrelevance of Private Concepts

- Some have gone further:

> Suppose everyone had a box with something in it: we call a 'beetle'. No one can look into anyone else's box, and everyone says he knows what a beetle is only by looking at *his* beetle. – Here it would be quite possible for everyone to have something different in his box. … But suppose the word 'beetle' had a use in these people's language? – If so it would not be used as the name of a thing. The thing in the box has no place in the language game at all; not even as a *something*: for the box might even be empty. (Wittgenstein 1967: §293)

- The 'beetle' here may be the internal concept: Wittgestein seems to be arguing that coordinated use is all that matters for meaning, and the concept itself is irrelevant.
- What can the internalist say in response?
  - Incidentally, is the referentialist immune to this argument?

The Nature of Concepts

Use and Reference for the Internalist

- That each speaker has their own internal language, or **idiolect**, and that what we see as the natural language English is just emergent from a bunch of very similar and causally related idiolects is surprising.
- It also prompts Frege's worry, one we can put pithily: **doesn't this make words and sentences about us, not the world?**
  - Or maybe: it makes them about the wrong bit of the world: *dog* is about DOG, a part of my mind or even brain, not canine animals.
- But there is a fairly immediate response: > when people utter a word whose meaning is a particular internal concept, they are not attempting to draw attention to the concept itself, but rather to things that … *fall under* that concept. (Elbourne 2011: 25)
- That is, a **convention** of deferred reference has arisen, so sentences activate concepts in hearers/readers, and those concepts are about the world, because things **fall under** or **satisfy** the concept.
- So what is 'falling under' a concept?

Concepts and Meaning

- The most obvious answer: a concept is or determines a **rule for classification**.
  - An object x falls under the concept CHAIR iff the rule associated with that concept says 'yes' to x.
- This seems to make a concept itself an entity with a meaning! For if a concept has a rule associated with it, that is equivalent to **assigning an intension** to the concept.
  - An intension I is a set of actual and possible entities. A yes-no classificatory rule R associated with a concept determines such a set: $I = \{x : R(x) = \text{'yes'}\}$; and vice versa: $R(x) = \text{'yes'}$ iff $x \in I$.
- This might appear to suggest a possible economy: why not think of a concept as a disposition to use a meaningful word, rather than being an independently meaningful entity?

The Causal Argument For Contentful Concepts

- One answer is that we need contentful internal concepts anyway, to understand **rational behaviour**.
  (4) Mental states causally explain behaviour in virtue of their contents.

  > We behave the way we do because of what we want and what we believe, and this seems to be just another way of saying that we behave as we do because of the contents of our beliefs and desires. (Brown 2022: §3.1; see Frege 1918: 310–11)

  (5) Causal powers depend on intrinsic features of the cause, i.e., the causal powers of a mental state are intrinsic to it. (Fodor 1987: 44)
  (6) The content of a mental state is intrinsic to it. (From 4, 5)
- The upshot: concepts are intrinsically contentful, since **they** cause my behaviour, not the things which fall under them.
- Moreover, if language use is rational behaviour – as when we say 'p' to reflect our belief that p – then that belief state is intrinsically contentful, not dependent (as the referentialist would have it) on the environment.

Concepts and Classifications

- If a concept is a rule, what sort of rule is it?
- It could be an tacit definition, some necessary and sufficient conditions we unconsciously apply.
- But the experimental data on classification doesn't support that – the **typicality effect** is the phenomenon where subjects are slower to judge that an unusual F falls under the concept than a typical one:

> The idea is that sparrows and wrens in some way correspond more closely to the concept BIRD than emus and penguins do, even though the latter do indeed fall under this concept. (Elbourne 2011: 27)

  - The tacit definition theory predicts no difference – if you are running through an internal checklist of conditions that a and b both satisfy, why should it take longer to verify one than the other?

The Prototype Theory

- The **prototype theory** says that a concept is instead a

> representation of features that the things [falling under it] can have, together with weightings indicating how important it is to have those features. (Elbourne 2011: 27)

- So for the concept BIRD , we might weight having feathers highly, flying quite highly, having wings highly, and then a bunch of other features lower.
  - In effect, we apply the concept BIRD to those things which closely enough resemble an exemplar bird in those respects which are bird-relevant.
  - This is **not** a definition: e.g., because it is **typical**, flying is on there, despite being neither necessary nor sufficient.

Problems for Prototype Theory

- The concept PET has a certain prototype: a dog, for example, is the canonical pet. A hamster, less so; a snake, even less.
- The concept FISH also has a prototype: a tuna? a mackerel? perhaps.
- What about the concept PET FISH? It has a prototype: a goldfish, obviously.
- The problem now is that the meaning of *pet fish*, which is the concept PET FISH, doesn't seem to be a straightforward consequence of the meanings of *pet* and *fish*, on the prototype theory.
  - For to fall under PET FISH is to be close to the prototype fish, and the prototype pet, and a goldfish is neither. (Fodor 1998: 100–108)
- Let's look at this in more detail.

Productivity and Compositionality

> One generates the productivity problem by asking how a finite creature could have a certain infinite epistemic capacity: how there could be infinitely many concepts that it can entertain. … a finite creature can get into an epistemic relation to an infinite set only by being in some epistemic relation to a finite object that specifies the set [suggesting] that each mental representation must itself by finitely specifiable, that the primitive basis from which complex MRs are constructed must be finite, and so forth. …
>
> Principle P provides a rough formulation:
>
> > P The interpretation … assign[ed] to a certain MR must be computed from the structural description assign[ed] to that MR.
>
> … presumably, [interpretation] assigns to the mental representation BROWN COW the intersection of the set of brown things with the set of cows. However, P further requires that [it] does so because … BROWN COW [has] a structure which includes the constituent representations BROWN and COW…. the structure of the interpretation… derives from the structure that [composition] assigns.(Fodor and Lepore 1996: 256–57)

- As we'll see later, this **compositionality** requirement is essential to systematic theories of meaning (lecture 7).

Compositional Concepts Aren't Prototypical: Missing Prototypes

> for indefinitely many Boolean concepts, there isn't any prototype even though: (1) the primitive constituent concepts all have prototypes; and (2) the complex concept itself has definite truth conditions. So, for example, consider the predicate "isn't a cat"; … there is a definite semantic interpretation for "is not a cat"; that is, it expresses the property of not being a cat, …. However, although "isn't a cat" is entirely well behaved on these assumptions, it pretty clearly has no stereotype; and nor do indefinitely many other Boolean complex concepts. (Fodor and Lepore 1996: 260)

Compositional Concepts Aren't Prototypical: the PET FISH problem

> an object's similarity to the prototype for a complex concept seems not to vary systematically as a function of its similarity to the prototypes of the constituents concepts. So, for example, a goldfish is a poorish example of a fish, and a poorish example of a pet, but it's quite a good example of a pet fish.
>
> … according to prototype theory, to have a concept is to have its prototype together with a measure of the distance between the prototype and an arbitrary object in the domain of discourse…. however, the distance of an arbitrary object from the prototypical pet fish is not a function of its distance from the prototypical pet and its distance from the prototypical fish. In consequence, knowing that PET and FISH have the prototypes that they do does not permit one to predict that the prototypical pet fish is more like a goldfish than like a trout or a herring, on the one hand, or a dog or a cat, on the other. But if prototypes aren't compositional, then, to put it mildly, the identification of concepts with prototypes can't explain why concepts are productive. (Fodor and Lepore 1996: 262–63)

The Bottom Line

- Fodor's NOT A CAT and PET FISH examples pose a significant challenge to the prototype theory.
- The argument ultimately rests on a **mismatch** between the logical structure of concepts, needed for compositionality, and the classificatory structure of concepts, needed to explain typicality effects (Fodor and Lepore 1996: 267).
- Fodor is no enemy of concepts however: he thinks they are necessary for any theory of mind. So while pessimism about prototypes may be warranted, he is not pessimistic that there are representational entities in the mind that play the concept role.
- So the fate of internalism should be separated from the fate of prototypes, without diminishing the fact that a task faces the internalist to give an account of concepts that allows them to play a role in cognition and word meaning.
- Or we could neatly sidestep the issue; referentialists don't need word meanings to predict typicality effects, and so don't face the problem of finding word meanings that can be compositional and prototypical.

References

References

Armstrong, D M (1989) *Universals: An Opinionated Introduction*. Westview.

Brown, Curtis (2022) 'Narrow Mental Content', in Edward N Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/content-narrow/.

Chomsky, Noam (1982) *The Generative Enterprise*. Foris Publications.

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

Elbourne, Paul (2011) *Meaning: A Slim Guide to Semantics*. Oxford University Press.

Fodor, Jerry A (1987) *Psychosemantics*. MIT Press.

Fodor, Jerry A (1998) *Concepts*. Oxford University Press.

Fodor, Jerry and Ernest Lepore (1996) 'The Red Herring and the Pet Fish: Why Concepts Still Can't Be Prototypes', *Cognition* **58**: 253–70. doi:10.1016/0010-0277(95)00694-x.

References (cont.)

Frege, Gottlob (1918/1956) 'The Thought: A Logical Inquiry', *Mind* **65**: 289–311. doi:10.2307/2251513.

Lewis, David (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy* **61**: 343–77. doi:10.1080/00048408312341131.

Mill, John Stuart (1882) *A System of Logic, Ratiocinative and Inductive*, 8th edition. Harper & Brothers. https://www.gutenberg.org/ebooks/27942.

Russell, Bertrand (1903) *The Principles of Mathematics*. Cambridge University Press.

Speaks, Jeff (2021) 'Theories of Meaning', in Edward N Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/meaning/.

Williamson, Timothy (2006) 'Conceptual Truth', *Proceedings of the Aristotelian Society* **80**: 1–41. doi:10.1111/j.1467-8349.2006.00136.x.

References (cont.)

Wittgenstein, Ludwig (1967) *Philosophical Investigations*, 3rd edition, G E M Anscombe and Rush Rhees, eds.; G E M Anscombe, trans. Blackwell.