Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Meanings Without Concepts

Florin Cojocariu

florin.cojocariu@s.unibuc.ro

University of Bucharest

Research Article

Keywords:

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11556097/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Meanings Without Concepts

Abstract

Simonelli's "Sapience without sentience: An inferentialist approach to LLMs" argues that mastery of the inferential role suffices for concept possession, so that large language models trained only on text can in principle possess all concepts. This reply accepts his semantics as a theory of meaning and denies it is a theory of concepts. Meanings tie words to words; concepts tie words to things. Simonelli passes from meanings to concepts in two steps: a rule describing seeing replaces seeing, and having the conceptual content of "red" is taken for having the concept red. In a variant of Jackson's experiment, an ordinary Mary with working color vision and partial mastery of "red" leaves her black-and-white room and cannot tell which tomatoes are red, although every premise of that rule is true of her. Frege and Wittgenstein kept concept and meaning apart but later on Sellars, Brandom and Simonelli brought them together. In the framework proposed here a word has two modes: one used of present things, part of pre-verbal patterns without content, and one used of absent things, placed among other words. A concept can be understood then as a link of the two modes. Mary lacks this link: its absence shows in her failed applications of "red", and others can correct them, so it is neither ineffable nor a matter of sentience. Language models are in Mary's position: they hold meanings more fully than any speaker and lack concepts that allow for words application.

Keywords: inferentialism, concept possession, theory of meaning, large language models, perceptual concepts, language-entry transitions

Introduction

This reply is somewhat unusual, because it agrees with the aims of the paper it replies to while it questions its definitions. I have two main points of disagreement with Simonelli's "Sapience Without Sentience" (Simonelli, 2026). The first is that inferentialism accounts for meanings and not for concepts. For me, a concept like red ties a word to the world, while a meaning ties words to other words. Simonelli's argument passes from meanings to concepts in two steps that support each other: it takes having a concept to be grasping its content, and it keeps the experiential dimension of a concept like red only as a sentence that describes it. The second point of disagreement concerns theoretical responsibility he invokes in the second half of his paper, which I find impossible to attach to a language model, mostly because there

is no bearer to attach it to. The question of the bearer in LLMs needs a paper of its own, so here I deal only with the first.

To argue about whether one has or does not have a concept, a minimum definition of "concept" should be provided, yet Simonelli does not provide one. Philosophically, "concept" behaves much like "being": its definition is almost impossible to agree on, and a short history of the word shows some of the reasons for the disagreement. This made it harder than expected for me to show that concept and meaning are different things.

Jackson's Mary (Jackson, 1982) helped, in a slightly different setup. In my version she is an ordinary person and not a colour scientist, and she masters only some of the inferential roles of "red", as we all do. On Simonelli's account she has the concept red inside the room. Yet once outside, in a garden of tomatoes of every ripeness, she cannot tell by sight which tomatoes are red, although her colour vision works. This makes one ask whether it is possible to have a concept one cannot use. It is the first hint that global strong inferentialism may miss something.

Language models, on the other hand, are truly hyper-inferentialist. Everything they receive, pictures included, reaches them as tokens, and every step they take goes from tokens to tokens. Yet a multimodal model can say which tomato in a picture is red, and Mary in the garden cannot. The question becomes what the model has that Mary does not.

This question leads to the main claim. Concepts and meanings are different animals that look identical only in the dark, and this paper tries to shed some light on them. Language models have a special kind of meaning for words, more complete than any speaker's, and they lack the concepts of words like "red", whose application we verify by seeing. Mary in the room is in the same position. She acquires the concept when she is shown red things, applies the word to them and is corrected.

A short history of the two words shows that, at least for Frege and Wittgenstein, meanings and concepts were different things. The identification of concept with meaning begins with Sellars, is established by Brandom, and is completed in Simonelli's semantics. The formalism of a framework developed elsewhere (Author, 2026) (referred in this version as "the framework" for review purposes) then lets us state precisely where the difference lies. It also shows why the difference disappears for theoretical words like "entropy", where the concept overlaps with the meaning and the model seems to have concepts while it has only meanings. The last section imagines what Sellars would have replied to this paper, and the reply tries to pre-empt some possible objections.

1 The problem with Mary

1.1 How Simonelli uses Jackson's experiment

Simonelli states at the outset what having a concept[1] is: "to grasp a concept is to master the inferential role of a linguistic expression" (Simonelli, 2026, p. 2). For a word like "red" this raises a difficulty, which he states himself: "one essential aspect of the meaning of 'red' is that it can be properly deployed, not only as the conclusion

[1] At first, this seems like a definition of concept, but it is not. There is something more about "grasping", which is typically used in "grasping a meaning", we'll discuss the confusion between concept and meaning at length.

of an inference (for instance, from 'scarlet'), but also non-inferentially, upon seeing something red!" (Simonelli, 2026, p. 5). The standard answer in inferentialism is to add language-entry rules, which connect perceptual circumstances to sentences.

Simonelli refuses them and gives a formal way of dispensing with the standard quasi-inferential rules. It answers the question "How can the fact that 'red' can be deployed non-inferentially in perception be captured in purely inferential terms?" (Simonelli, 2026, p. 6). The answer makes a description of seeing stand in for seeing:

$$red(a),\ color\ vision(S),\ looking\ at(S,a),\ good\ lighting \vdash sees(S, \langle red(a) \rangle)$$

I will call this the S-rule. Its conclusion is the description of an experience of seeing, not the experience itself.

Simonelli concludes that "inferences, properly so-called, suffice to completely articulate the *conceptual content* of even a term like red, which is essentially such as to be deployed in perception" (Simonelli, 2026, pp. 6–7).[2]

Building on this, Mary the colour scientist is said to have "reached the theoretical limit of what can be known about the colors without actually having experienced them" (Simonelli, 2026, p. 7). Since she knows all the ways in which "red" can be used, it follows that "she knows what it is to see that something is red" (Simonelli, 2026, p. 8). This gave me pause again, because for me it is obvious that she does not know what it is to see that something is red, but only knows how seeing that something is red can be described.

He then turns to the sticky problem of what Mary lacks. Here the argument is simple. "The question is: what elements of the *conceptual content* of 'red' is she lacking *understanding*?" (Simonelli, 2026, p. 8). Anything that can be said in answer can be understood in inferential terms, and Mary grasps it. Whoever insists that she lacks something is left saying that it is "inarticulable, indeed, ineffable" (Simonelli, 2026, p. 8). But for the inferentialist "every bit of the conceptual content of every expression is effable" (Simonelli, 2026, p. 8). So Mary lacks no *conceptual content*. She lacks only the experience of seeing red, and this, Simonelli notes, both parties acknowledge.

From this he concludes that one can "possess a concept that has an essentially experiential aspect without actually having the relevant experiences" (Simonelli, 2026, pp. 8–9). The Vulcan psychologist, who talks with full understanding about grief, joy and guilt without ever having felt them, is a second case. Both have "a concept which is essentially such that it can be deployed non-inferentially in cases in which one has certain experiences, without actually having the relevant experiences" (Simonelli, 2026, p. 9). This is the step that opens the way to language models, which would have such concepts without any experiences at all.

1.2 Two claims with problems

Simonelli's argument thus makes two claims:

- Having all the inferential rules for "red" amounts to having the concept red, and with it to knowing what it is to see that something is red.

[2] original emphasis removed, emphasis added; my emphasis everywhere from here on

- Whatever Mary is said to lack is either sayable, and then she grasps it inferentially, or ineffable, and then it is not *conceptual content*. So she has the concept.

Both seem false to me. I think the real problem is that the totality of the descriptions of experiences involving a concept is taken to be the concept itself. This happens in two steps, and the transition between them is not explicit. In the first, experiences are taken to be the same thing as descriptions of experiences. In the second, having the *conceptual content* of a concept is taken to be the same thing as having the concept. The S rule achieves the first. The Mary argument achieves the second, in a hidden way.

The S rule *makes perception part of the content*, although it is hard to see why seeing or feeling would be content. If seeing and feeling are not content, it is odd to accept descriptions of seeing as valid inferential roles for actual seeing without any difficulty. Put plainly, the S rule takes descriptions of (non-conceptual) sentience to be (inferential) sapience.

The second step is hidden in the Mary argument. The experience Mary lacks is set aside as unimportant, and the question is moved to the *conceptual content* of "red". The argument concludes that she lacks no *conceptual content*, and on the same page that she possesses the concept (Simonelli, 2026, p. 8). Nothing connects the two conclusions except the assumption that having the conceptual content is having the concept. The same assumption is visible where Simonelli describes the theory he wants, Dummett's "*full-blooded*" theory of meaning, "*a theory that aims to give a non-circular account of the concepts expressed by the primitive expressions of a language*" (Simonelli, 2026, p. 6): a theory of meaning is taken to be an account of concepts.

The first step is what makes the experience unimportant. Once the S rule has put the description of seeing into the content of "red", Mary already has everything about seeing that counts, and lacking the experience costs her nothing conceptual. It is the same move in both places: what can be said about seeing is taken for what seeing gives.

To better grasp this, let us leave "red" for a moment and take a more difficult case, "love" and "being in love". Most of us love or have loved somebody, and we describe this with the word "love". But before it happened, we had been told what love is by innumerable poems, novels, films and weary adults. We knew love by deference.

The truth is that nobody really gets it until it happens to them. Someone who knew, in theory, all the inferential roles of "love", as an LLM probably does (more on this later), would still not be prepared for the real experience. The claim that such a person knows what it is to be in love is difficult to sustain. Simonelli sustains it all the same: an LLM "can know what it is to give a damn without itself giving a damn" (Simonelli, 2026, p. 16).

Simonelli would say that what such a person lacks is the experience of love, not anything conceptual. Yet someone who has been in love, with few of the inferential rules, knows what it is to be in love, and someone who knows all of them and has never been in love does not *know what it is to be in love*.

Or, the better question: who has the fuller concept of love, the real person having few of the inferential rules but having experienced being in love, or the ideal person knowing all inferential rules but never being in love for real? To me, the answer is undeniable: *we do possess the concept of love if we have ever been in love*. This raises

the question whether possessing a concept is only a matter of possessing inferential roles. And also a subtler one, if "experience" can be assimilated with "sentience".

Inferentialism develops from Wittgenstein's famous claim that meaning is use, with inferentialism being a way to formalize what "use" is. There is much to say about how Wittgenstein relates to Simonelli's paper, including the use of *Philosophical Investigations* (PI)[3] §261 in his footnote 11, which seems to me to take the remark further than Wittgenstein does.[4] These points are secondary. More important for my response is that Simonelli takes his purely inferential version of global strong inferentialism to be the one actually defended by Sellars and Brandom (Simonelli, 2026, p. 6). I do not plan to enter that argument, because it would easily overflow the boundaries of a single paper. I keep instead the idea that both Sellars and Brandom distinguish intralinguistic moves, language-entry transitions from perceptual circumstances, and language-departure transitions into action. What matters for my argument is that inferentialism does need entry and exit transitions, the quasi-inferential rules, for a very good reason. Explaining this will be the bulk of the next sections.

The paper's argument needs the hyper-inferentialist position to be true, and it needs to identify a concept with mastery of the inferential roles of the word that names it. To show why both are problematic, we need a different Mary experiment.

2 Normal Mary experiment

2.1 A slightly different setup

Jackson used his original mental experiment (Jackson, 1982) against physicalism: if Mary the color scientist knew all the physical facts and still learned something else on release, not all facts are physical.

Simonelli replaces complete physical knowledge with complete inferential mastery "*[..] she knows just what it is for something to be red, since she grasps all of the inferential relations between sentences that articulate the content of 'red.'*" (Simonelli, 2026, pp. 7–8) He does not bring her outside to experience red and green, because he does not need to.

---

[3] I quote PI in two English translations. §§40 and 43 are quoted from Anscombe's translation (Wittgenstein, 1953/1968), through which "the meaning of a word is its use in the language" entered English-language philosophy. All other remarks, PPF §365 included, are quoted from the revised translation by Anscombe, Hacker and Schulte (Wittgenstein, 1953/2009). Both editions print the German text facing the English. I use both because the two translations differ at a point this paper depends on. Anscombe renders both "Gebrauch" and "Anwendung" as "use", so that in her English Wittgenstein's remarks on the meaning of a word and his remarks on the analysis of a concept are both about "the use of a word". Hacker and Schulte keep "use" for "Gebrauch" and give "application" for "Anwendung". The German keeps the two apart, and Anscombe's English joins them; the clearest case is PI §383, discussed below.

[4] Simonelli quotes PI §261 (Wittgenstein, 1953/2009) to describe the position of someone who insists that Mary lacks something: asked what she lacks, such a person "would like just to emit an inarticulate sound" (Simonelli, 2026, p. 8). In his footnote 11 he notes that the remark belongs to the private-language argument and sets that argument aside, taking inferentialism as a premise (Simonelli, 2026, p. 8, n. 11). The illustration still borrows more from the remark than the remark contains. Wittgenstein is discussing the attempted private designation of a sensation by the sign "S", and the sentence that follows Simonelli's quotation reads: "But such a sound is an expression only as it occurs in a particular language-game, which should now be described" (Wittgenstein, 1953/2009). The inarticulate sound is sent back to a language game to be described. It is not left as something ineffable. Wittgenstein's question there is how an expression acquires meaningful use, and §261 says nothing about whether every aspect of experience is captured by linguistic descriptions. Section 7 argues that what Mary lacks is of the first kind: it belongs to the practice in which words are applied to things.

Compared to Jackson's version, we'll change only what Mary is and knows about colors: she is a normal person, not a scientist, and color theory may be foreign to her. She lived her entire life in a black and white controlled environment and she masters only a limited number of inferential roles for "red", as we all do.

One day she is brought out to a garden and asked to pick only the red tomatoes, while leaving the green ones to ripen. Before her release she had rules of using "red" and "green" but never applied them to anything that *she saw as red or green*. Now she sees the tomatoes and their different colours, but cannot tell which is red.[5]

2.2 Can we have concepts and not be able to recognize what they refer to?

Simonelli claims that Mary has the red and green concepts because of hyper-inferentialism (or Global Strong Inferentialism): mastering the inferences that a word figures in suffices for her having the concepts. He explicitly rejects the language-entry rules (or quasi-inferences), otherwise the hyper-inferentialist claim will fall (pp. 5-6). This is an elegant solution to the problem of circularity and primitive expressions, but it creates a new problem. Mary has, in this account, possession of the concept of red, although she cannot use it in practice. Nothing in principle prevents people with total achromatopsia from having the concept red (p. 8), and he even suggests that Mary is more realistically imagined as completely color blind (p. 7, n. 9). But normal Mary is not color blind: she sees the colors and still cannot tell which one is red, so her failure cannot be put on perception.

And this is not about sentience or sapience, our Mary possesses both of them when she cannot tell which are the tomatoes she is supposed to pick. She is sentient even by Simonelli's own definition, where perception is the receptive aspect of the perception/action nexus through which a creature navigates its environment (p. 9). She has inferences with red including the S rules, and she cannot tell which colour is red. In fact, normal Mary is a counterexample to the S rule itself: all its premises are true of her, the tomato is red, she has color vision, she is looking at it in good lighting, and still she does not see that the tomato is red. Simonelli notes that the rule is defeasible (p. 6, n. 7), but the premise needed to save it, something like "S has learned to apply 'red' to what S sees", is a description of precisely the language-entry transition he removed: Mary can know this premise, but it is not true of her, since she has never learned to apply "red" to what she sees. Whether knowing it without it being true of her amounts to having the concept is the question of section 4. Even if she had all inferences (as Simonelli seems to suggest in his version) this would not change. It appears that having a concept is something other than mastering the inferential role of the words naming them. Moreover, the fact that someone can teach her what precisely is red in the garden (or train as Wittgenstein will say), shows that concepts have to do with application and not only with inference.

[5] Congenitally blind speakers command the inferential roles of colors in the same way as Mary, and there are cases of people gaining vision for the first time at a later age and experiencing the same problem as Mary, only at a much bigger scale: right after the operation they cannot match the shapes they see with the shapes they know by touch, and they learn it in a few days, by applying their new vision to things (Held et al., 2011).

Before moving to our main claim that shows what is really missing for normal Mary and what she really has from the inferential role mastery, language models show us an apparent problem.

3 A language model is hyper-inferentialist

Simonelli's thesis is about pure LLMs, trained on nothing but text (p. 12), and such a model cannot even be asked about pictures. The obvious objection comes from multimodal models.

Suppose now that we show a multimodal language model a picture of a red tomato and one of a green tomato and ask it which is which: it will have no problem doing so. What is happening here, why can an LLM tell the colors with no real sensing experience?

The short answer is that the model does not see anything, it is handed a token representation of a tomato: the picture is cut into patches and an encoder turns each patch into a vector in the same space as the word tokens, and the association with color names is already carried by those vectors because it was formed in training. The picture is, for a trained multimodal model, already an encoded representation as every sentence is.[6]

One may reply that human perception is encoded in the same way, but such a reply describes the human brain as something it is not. The retina is brain neurons (albeit specialized ones), connected directly to other brain neurons. In seeing they are exposed to light and generate electrical impulses in an immensely complex biological neural network. In a human the seeing can happen with no words involved at all, while in a model "seeing" is only tokens. Encoding assumes a code and nothing of the sort can be found there (see Brette, 2019). In an LLM the encoding is in the translation rules from picture to tokens and then in the associations between tokens formed in training by humans that know what is red or green.

Moreover, there is no exit point either, a model only outputs text, it has no need to act in any other way. It is because of that that a model is truly hyper-inferentialist and Simonelli's semantics gets the model precisely right. And it is also true that if the only test is a semantic one, a human and a model cannot be told apart; the only difference appears with the entry and exit points. Simonelli agrees that pure LLMs lack the perception/action nexus (p. 13), but he puts this difference in sentience; normal Mary shows that it belongs to concepts. Properly speaking, an LLM does not recognize anything, it just outputs text that *we see* as standing for a recognition event.

4 Meaning and concept are different things

Here is the argument in a nutshell: meaning lives entirely in language and it is inferential. It is meaning and not concept that Wittgenstein attached to a word use. Concept is a different entity and its use and its history all point to the fact that concept has to do with the application of language in the real world.

[6] For the model there is no difference between sentences and pictures, both are just string of tokens it receives.

Ordinary grammar already marks the difference. We speak of the meaning of the *word* "horse" and of the concept horse, and not of the concept of the word "horse." "The meaning of horse" is heard as shorthand: the word is understood as the subject. "The concept of the word horse" is not shorthand for anything, it is simply not meaningful.

If we replace "concept" with "meaning" in Simonelli's paper, it will be difficult to object to it because having a meaning of a word is almost by definition having its inferential roles; no entry and exit points are required for having a meaning. And our Normal Mary, in the garden, has the meaning of "red" and "green", and you do not need meanings to apply words to reality; it is the concepts that you need.

Arguing for this difference is, though, more complicated. We'll look first at how we do use concepts and meanings in the real world, and then, with our normal Mary, at whether we need concepts to refer to things. We'll then see why what LLMs have is meanings and not concepts, and finally turn to Peacocke for a different view on having a concept.

4.1 We use concept and meaning differently

First, we adopt here Wittgenstein's view in PI §43 (Wittgenstein, 1953/1968): a meaning of a word in a language game, in a practice is its use. It is public, settled and corrected only by the practice of speakers, and can be recorded in textual descriptions.

The practice is what people do, and the meaning is what the practice settles in language. The practice is correcting, going on, withholding assent; it takes time, it can lapse, and people join and abandon it. The meaning is what that activity keeps in place in a language: which uses are licensed and what commits one to what when using a word. It results that one can hold a meaning without a practical application in the real world, and most of our learning is based on that.

While meanings are also public (because established and corrected in a practice), concepts are only personal, something that can be seen even in the way we use the word: we speak about "someone's conception of something" as something subjective and personal. This comes from the fact that we build our own concepts about the things and the world in our own interaction with the things and the world (a model for this view is detailed in the next sections). They are still corrected in public, but by things and by people pointing at things ("that one is not red yet, it is pink"), while meanings are corrected by other sentences. This is why we can speak of a concept of red that is ours, shared by those who were corrected in the same way. Concepts are primarily devices to interact with reality, and using them can run without descriptions or sentences altogether.

There are two reasons why is difficult to see this: The first is that while concept and meaning are different, they change and evolve together. The second is because of what we can call "theoretical concepts", new terms that are defined inside theories and are justified exclusively by inference inside said theories. A concept like red is always applied to experience and corrected by it in case of misuse (taking pink to be red makes you pick a tasteless tomato). But a concept like "spin" in particle physics is one that cannot be verified directly in experience. There is measurement, but that is something very far removed from the real particle, effects many orders of magnitude larger that are *first established in theory.* The difference is in what corrects the concept:

red is corrected by the things themselves, spin only through effects that a theory says it has. The fact that what we call theoretical concepts are in fact meanings of words in a specific practice and language game in science, will be argued once we have the proper framework to clearly disambiguate all these notions in section 7.

We pick the red tomatoes in the garden and leave the green ones, usually without saying or thinking a sentence. The concept red is used in the picking but does not require a sentence in order to be applied. We know the tomato is "red" and can talk about it at once, but nothing is said in the picking itself. Picking ripe tomatoes is not an inference through the meaning of "red." I do not go from "this is red" to "so it is ripe" to "so I pick it" and then I pick it. What the picking exercises is the tie between the word and what I take in, but it still stays answerable to the word: "that one is not red yet, it is pink" said by someone, corrects it.

It is important to see that the picking is still not word-free. A bird picks red berries with no word at all, but a human never picks anything as a bird does. The word "tomato" was learned in the language game as part of what the child takes in of tomatoes, at first as a motor pattern[7], and it is at hand when she sees one; she may hear it said in her mind when seeing a tomato when she also expects the smell of its leaves. A word at hand in this way is applied, takes part in such expectations, and infers nothing. Inference runs on the meaning, and the meaning is idle in the picking, because there is no inference needed.

For "red" the word is at hand in the same way: learned together with the color that red things share, and ready to be said of this tomato or that car. But contrary to meaning, having the concept of red and using it can be the difference between life and death: telling an edible mushroom from a poisonous one, stopping at a red light, keeping a hand away from a very hot iron. When you brake at a red light, you do not say to yourself "I need to brake because red means I need to stop." What you learned and use in braking is the concept of a red light. The meaning of "red light" is what you use when you explain *in language* afterwards why you braked.

So, regarding the way we use the words, we established that while meanings live in text and are public in the sense that are corrected in a public practice, concepts are devices for interaction with the world and are subjective and personal, verified in experience. While they are indeed connected and evolve together, they seem to have a very different nature.

4.2 Do we need concepts to refer to things?

Mary, before she sees color, has the meaning of "red" and not yet the concept. Her word "red" still refers to real red things, because she is inside the practice even in her black and white world, answerable to and deferring to those who apply the word by sight - the division of linguistic labour, (Putnam, 1975).

Through them her "red" has its reference in full, even if her own mastery of its inferential roles is limited. The concept comes when she can apply the word herself: when she is shown red things, applies the word to them and is corrected. Until then she just infers that something is red, in the way physicists infer a particle has a certain

[7]This will be fully explained in section 7

spin from a measurement: she sees the gray and infers the red, as physicists see the reading and infer the spin.

Mary's "red" refers to red things, and she has no concept red. Reference does not need the concept, it operates on whatever those who apply the word by sight established to be red ("tomatoes are red and tomatoes look like this - gray for Mary"). But once in the garden, Mary's use of red goes further than borrowed reference. She can ask someone which tomatoes are red, pick the ones she is shown, and be corrected by the next picking. Her borrowed applications end in her own action on the things. The model's answers end in text, and any action on things is taken by someone who reads them. One can argue that if Mary did this in the room, she would couple "red" to a shade of gray she takes in and would have a concept red different from ours, as Wittgenstein's color-blind tribe (Wittgenstein, 1977, I §13) has color concepts different from ours. But this becomes problematic because it overlaps with the concept of other colors and a certain gray, all producing the same narrow band of shades of gray that red appears to be in black and white. This will lead her to be unable to tell a green tomato from a red one when they have a similar shade of gray. Her inferences, like $red(a),\ green(a) \vdash \bot$, will tell her that something is wrong, but not which tomato is red.

A language model is in Mary's position before the garden: it has the meaning of "red" through those who wrote the texts it was trained on, and its "red" refers to red things through the same people Mary depends on, those who apply the word by sight and trained the model. Unlike Mary, it has no gray to take in and obviously cannot pick any tomatoes; its borrowed applications never end in its own action on the things. It has the meaning of "red", very likely a more complete one than Mary's or ours, but it has no way to come to the concept.

4.2.1 LLMs have the meaning of words more completely than any speaker

Simonelli notes in a footnote that "It will turn out to be a consequence of the account to follow that, beyond just possessing concepts, an LLM can in fact possess a much fuller understanding of any material concept than any of us might possess." (p. 5, n. 4) I agree, with one correction: what models have more completely is meanings. In my reading LLMs can never have the concepts of words like "red", whose application we verify by seeing[8].

Each speaker learned a word from particular people and holds a partial version of its public use. A model trained on a corpus holds the public use as far as text records it, and goes on from it to new cases; in this sense it holds far more of the public use than any speaker.

From the moment Wittgenstein introduced the notion until LLMs appeared, "the public use" was an elusive object. It is what LLMs now have encoded in their weights. A language model has no use for concepts in the sense discussed here. Its only output is text, and nothing it produces is applied to things by it or checked against them for it.

[8]They can have what we call "theoretical concepts" like "spin", but we'll see below that in my view these are meanings rather than concepts.

Training, including the preference training Simonelli describes (p. 12) [9], does not supply the entry and exit points either because it is a loss computed over a distribution; it shapes what the model holds and addresses no one's experience.

4.3 On Peacocke's account Mary does not possess the concept red

Peacocke individuates a concept by its possession condition: what a thinker must satisfy to have it, stated without presupposing that she already has it (Peacocke, 1992, Section 1.4). For an observational concept like red, the possession condition involves perceptual experience. The thinker must be disposed to judge that something is red when experience presents it as red, taking the experience at face value.

Mary in the room, whether she has every inference for "red," as Simonelli's Mary does, or only some of them, as ours does, has no such experience. She meets Peacocke's attribution conditions, under which beliefs involving red can be ascribed to her through deference, but she does not meet his possession conditions.

Peacocke's account is a theory of concepts and was not written against hyper-inferentialism, yet it states the requirement this section argues for: a concept like red cannot be had without what is taken in of red things. It also gives a reason why the requirement cannot be met inside language. If the possession condition is not to be circular, what the concept is anchored in must be available before the concept.

Peacocke meets this requirement with non-conceptual content: experience represents the world as being a certain way, accurately or inaccurately, without the subject needing the concepts that describe it. I do not take over this part of his view: section 7 keeps his requirement and offers a pre-verbal level that is not content.

We can conclude this section by saying that there are a number of strong reasons to differentiate meaning from concept and to connect the first to the language use (inference roles of words) and the second to the application of a word to our perceptions of the world.

A further argument is historical: when meaning and concepts start to be articulated in analytic philosophy at the end of the 19th century, they were treated as very different things.

5 Frege and Wittgenstein kept meaning and concept apart

5.1 Frege's views

Frege's German keeps the two on opposite sides. A concept-word has a sense (Sinn) and a reference (Bedeutung), and the reference of a concept-word is a concept.

In 1891 Husserl sent Frege his *Philosophie der Arithmetik*, and in his reply of 24 May Frege set out the distinction between sense and 'Bedeutung'[10] that he was then writing up in "Über Sinn und Bedeutung" (Frege, 1997, pp. 151–171). The letter

[9] RLHF, where human evaluators rank the model's outputs.

[10] We keep 'Bedeutung' untranslated because as we'll see the translations were part of the reason meaning and concept were conflated later on.

contains a schema with three columns, for the proposition, the proper name and the concept word: each sign expresses a sense, and through the sense it has a Bedeutung. For the concept word the Bedeutung is the concept, and from the concept there is one more step to the objects falling under it. Take "horse": the word expresses a sense, the way the concept is given to anyone who understands the word; through the sense it refers to the concept horse, which *for Frege* is a function giving the True for every horse and the False for everything else; and under the concept fall the individual horses, Bucephalus among them. Frege then draws a second schema, for what he takes to be Husserl's view, and there the concept word goes from its sense directly to the objects; the row where the concept was has disappeared. Simonelli's inferentialism is drawn on this second schema, with the concept on the side of the sense. *I do not take over Frege's account of what a concept is*, but his schema keeps the place my argument needs: the meaning of a word stands where Frege has the sense, and the concept has its own row, between the meaning and the things (Frege, 1997, pp. 149–150)

What belongs to the word is its sense. The concept is what the word stands for. It is on the side of the world, with the objects that fall under it: horses fall under the concept horse.

In his reply to Kerry, Frege takes this further, in the paradox that the concept horse is not a concept (Frege, 1997, pp. 184–185). A concept is predicative: it is what is said of objects, as in "Bucephalus is a horse." When we name it, to talk about it, as in "the concept horse," the name in subject position stands for an object, and the concept is no longer there. So whatever we say about a concept is said about something else, and the concept itself is present only when it is applied to a thing. *Talk about a concept does not contain the concept.*

This is Mary's situation. She has everything that can be said about red, all the sentences of the form "the concept red is...," and the S-rule is one of them. What she does not have is the concept at work, applied to this tomato.

Frege does not support every part of this paper's claim. His concept is objective and is had by no one, and his application is logical predication, not perception. What matters here is that he supports two claims: the concept does not stand with the sense of a word but between the sense and the things, and a concept is had in its application, not in what is said about it. The claim that a concept is had by someone, who applies it to what she sees and can be wrong, stays as formulated in the arguments of the previous sections.

5.2 Wittgenstein's views

Wittgenstein took Frege's word 'Bedeutung' and moved it to the word's side. In his use the Bedeutung of a word is no longer what it stands for; it is its use, so that it becomes what we call today 'meaning'.

PI §40 (Wittgenstein, 1953/1968) separates the meaning of a name from its bearer. When Mr N.N. dies, we say the bearer of the name died, and not that the meaning of the name died.

PI §43 (Wittgenstein, 1953/1968) then explains "Bedeutung" for a large class of cases, though not for all: "Die Bedeutung eines Wortes ist sein Gebrauch in der Sprache."

Later on, Wittgenstein's German leaves room for the distinction: concepts are tied to application and to facts, and meanings to use. But Anscombe translated 'Anwendung', which is rather 'application', as 'use', the word she used for 'Gebrauch'. Hacker and Schulte's translation of PI §383 has it right (Wittgenstein, 1953/2009):

> We do not analyse a phenomenon (for example, thinking) but a concept (for example, that of thinking), and hence the application of a word. So it may look as if what we were doing were nominalism. Nominalists make the mistake of interpreting all words as names, and so of not really describing their use, but only, so to speak, giving a paper draft on such a description.

Wittgenstein does not use the two words as terms of the distinction we made in this paper. What matters is that he sees two different standings for a word: its use in language (which is its meaning), and its application, which is what analyzing a concept examines.

On this distinction, the builders' game of PI §2 and the remark in PI §7 (Wittgenstein, 1953/2009) that a language game includes the actions into which language is woven describe both. The builder's bringing a slab is the application of "slab"; what a description of the game records is its use.

We can also find in PI §242 (Wittgenstein, 1953/2009) a relevant observation which connects directly to Mary's case and shows why when she gets outside she cannot tell red from green:

> It is not only agreement in definitions, but also (odd as it may sound) agreement in judgements that is required for communication by means of language. This seems to abolish logic, but does not do so. It is one thing to describe methods of measurement, and another to obtain and state results of measurement. But what we call "measuring" is in part determined by a certain constancy in results of measurement.

In the sense of this paragraph, agreement in definitions is agreement in meanings while agreement in judgments is agreement in applications, and it is precisely the last one that Mary lacks when she has to tell red from green.

There are a number of remarks scattered in different writing that show Wittgenstein seeing concept as application to world, while meaning as use in language. For example, in "Philosophy of Psychology, A Fragment", at a later date he writes in PPF §365 (Wittgenstein, 1953/2009):

> If concept formation can be explained by facts of nature, shouldn't we be interested, not in grammar, but rather in what is its basis in nature? – We are, indeed, also interested in the correspondence between concepts and very general facts of nature. (Such facts as mostly do not strike us because of their generality.) But our interest is not thereby thrown back on to these possible causes of concept formation; we are not doing natural science; nor yet natural history — since we can also invent fictitious natural history for our purposes.

Wittgenstein says here that our concepts are formed against very general facts of nature: imagine those facts different, and different concepts would look natural. A word's inferential role can include sentences stating those facts. But knowing the sentences is not the same as having concepts formed under the facts, and a language model has only the sentences, and Mary has only the sentences for colors.

In "On Certainty" §65 (Wittgenstein, 1969/1972) he writes:

> When language-games change, then there is a change in concepts, and with the concepts the meanings of words change. ("*Wenn sich die Sprachspiele ändern, ändern sich die Begriffe, und mit den Begriffen die Bedeutungen der Wörter.*")

Here Wittgenstein names both in one sentence and keeps them apart: concepts change with the practice, and meanings change with the concepts, a point we'll also make in section 7.

There are a number of remarks on rule-following that are relevant to us in the measure in which inference is a variety of rule-following in inferentialism. They all point to where the application happens: a rule never contains its own application, and adding rules only adds more that must be applied.

For reasons of space, I will not go through them[11], but their point is this: any course of action can be brought into accord with a rule, and equally into conflict with it. What settles the matter is *a grasp of the rule shown in following it and going against it*, that is, in being corrected.

Following a rule is a practice. In Simonelli's semantics an inferential role is specified by rules, sequents of a calculus, and by PI §201 (Wittgenstein, 1953/2009) no specification contains its own application. The role is the meaning, and the application is where the concept is at work.

6 The identification of concept with meaning begins with Sellars and is established by Brandom

In §29 of *Empiricism and the Philosophy of Mind* (Sellars, 1956/1997), Sellars holds that we are not aware that something is red before we have the word "red," and I agree with this. But a concept is what lets us take something as being of a kind, as red or as a tomato. So if every awareness of kinds is a matter of language, concepts are a matter of language too, and having a concept becomes knowing how to use a word. This is where concept and meaning start to be treated as the same thing. Sellars does not ignore the relation with the world. In §19 he says that we learn the concept green through a long history of habits of response to green objects in different circumstances, but that reacting with "This is green" is not enough to have it: we must also know that these are the right conditions for telling colors, so we have the concept only together with many others.

Sellars comes closer to the view of this paper than the passage on psychological nominalism suggests. In "Some Reflections on Language Games" he distinguishes three kinds of transition in a language: language-entry transitions, from perceptual circumstances to sentences; intralinguistic moves, from sentences to sentences; and language-departure transitions, from sentences to action (Sellars, 1954). Entry transitions are not obeyed rules. They are pattern-governed behaviour, acquired in training, and the speaker who says "this is red" of a red thing does not infer it. In his later work Sellars adds picturing, a relation between linguistic objects, considered in their natural characteristics, and the things in the world, which is neither semantic nor normative (Sellars, 1962). In my reading, Sellars has both modes of use of a word that section

[11] The so inclined can check PI §§198, 201, 202, 208, and 217.

7 distinguishes: the responsive "this is red," said of something present, and "red" in moves between sentences. What he does not have is the link between them. In his later work the functional role of a word, what his dot-quotes classify, includes the entry transitions together with the intralinguistic moves (Sellars, 1974), so the two uses become parts of one role. Picturing stays outside the role, in the natural order.

The encounters with things are where the concept comes from; the concept itself lives in the whole system of concepts. On Sellars's account, then, our normal Mary in the garden is expected not to be able to tell red from green, because she never had the history through which the concepts red and green are acquired. Brandom keeps the encounters in his account of language, as the way perception enters it. Simonelli drops even this, and his Mary has the concept with no history of encounters at all.

Brandom defines the conceptual as the inferentially articulated, and keeps language-entry and language-exit moves in the content of empirical concepts, even though, as he says, these moves are not themselves inferences (Brandom, 1994, Chapter 4).

For Brandom, reliable differential responses to red things are not conceptual. We share them with artifacts such as thermostats (Brandom, 1994, p. 5), and a parrot trained to say "red" when it sees red things has them too. What makes a human report conceptual is that it is inferentially articulated: the one who says "this is red" can also say what follows from it and what would count against it.

Brandom distinguishes three kinds of inferentialism. For the weak one, inferential articulation is necessary for conceptual content; for the strong one, broadly inferential articulation is sufficient; for hyper-inferentialism, narrowly inferential articulation is sufficient. He says he defends the strong version and not the hyper one (Brandom, 1994, pp. 131–132) [12]. So for Brandom the content of "red" is its broad inferential role, which includes the circumstances in which the word is properly applied without any inference.

Simonelli completes the identification. He rejects language-entry rules in the semantics, so that nothing is left of a concept but the narrowly inferential role of a word (Simonelli, 2023; 2026, Section 2).

Read in one line, the history of the word "concept" looks like this. Frege put the concept on the side of the world, as what a concept-word stands for, with the sense on the side of the word. Wittgenstein put meaning in the use of the word and kept concepts tied to its application. Sellars moved every awareness of kinds into language but kept the encounters with things as the history through which a concept is learned. Brandom kept those encounters in the content of empirical concepts, as circumstances of applying a word without inference. Simonelli removed them, and in his semantics the concept is the narrowly inferential role of a word. By the grammar test, this is its meaning, and Simonelli says as much: knowing the conditions under which "red" can be used non-inferentially "is all that is required in order to be counted as knowing the meaning of 'red'" (Simonelli, 2026, p. 8).

[12] The two divisions use "strong" differently. Simonelli divides inferentialism by whether mastery of inferential role is necessary or sufficient for having a concept, and by whether this holds for some concepts or for all; his own view is "global strong inferentialism" (Simonelli, 2026, pp. 3–4). Since the only rules he admits are narrowly inferential, his global strong inferentialism is what Brandom calls hyper-inferentialism. Admitting language-entry rules makes the view global weak inferentialism on Simonelli's division (Simonelli, 2026, p. 6), and strong inferentialism on Brandom's, which is the version Brandom says he holds. Simonelli argues that Brandom is nevertheless a hyper-inferentialist (Simonelli, 2023).

Read this way, his argument is sound. Training on text can deliver mastery of inferential roles, and so it can deliver meanings, which is what section 4 granted the model. The conclusion that models have concepts follows only if concepts are meanings, and the previous sections showed that the two come apart, both in Mary and in the model.

The disagreement is over more than a word. "LLMs possess concepts" can be read in two senses: that they hold the meanings of their words, or that they can apply their words to the things the words are about. The conclusion is heard in the second sense, and Simonelli's argument supports it only in the first. If he holds that Mary in the room has the concept red, he holds that having a concept does not include being able to tell, by seeing, which things fall under it.

The identification also leaves unexplained what normal Mary lacks in the garden. Once concept and meaning are separated, the puzzle has an answer, and the framework notation, which has symbols for the non-verbal part of the story, makes the answer simple to state.

7 The framework unifies the cognition of basic minds with language

Inferential notation has symbols only for sentences and the transitions between them. Mary's seeing, which she has, and the tie between her seeing and the word, which she lacks, are not sentences, so the notation cannot state either.

The framework, introduced by Author (2026) [13], has notation for the pre-verbal side of cognition and by this clarifies a series of problems and paradoxes at the interface of the language with the world. For the purposes of this reply, the framework can be treated as a meta-language that formalizes basic cognition (pre verbal cognition), disambiguating between the way we use words in relation to present objects and in their abstract mode. Here are the basic elements that will help us clarify that indeed concepts and meanings are different things.

7.1 Pattern-constellations

A pattern-constellation $\{X\}$ is the integrated grasp of a thing formed in repeated encounters that mattered to someone, as in $\{FIRE\}$.

- $\{X\}$ is pre-verbal. It can form before the word, it consists of no sentences, and animals have it.
- $\{X\}$ contains sensory, motor and feeling patterns, learned together, not one after another.
- $\{X\}$ is someone's: formed in her encounters, and brought to bear on what she meets.

$\{X\}$ is not a pure abstraction, in physical terms it is a pattern of recurrent activation across functional zones in a neural network (artificial or biological), formed by Hebbian learning, and it can be best modelled as an attractor in a Hopfield network

[13] A framework like this one is in continuous development, preprints are published as intermediary versions. The most up-to-date version can be consulted here URL removed for review and is highly recommended because significantly updated in both definitions and the applications.

(Hopfield, 1982). For example $\{FIRE\}$ is constituted by learned perceptions (light, heat, movement patterns, smell), reactions (close by to get warm, escape if too hot), feelings (warmth of the fire in winter, danger if in contact with). These are not separately stored, and then computed in some sort of inferences, but learned together and applied together when a fragment is available.

$\{X\}$ meets Peacocke's non-circularity requirement without being content (Peacocke, 1992, Chapter 1). Content is delimited by correctness conditions: it draws a line through the ways the world could be. Peacocke's scenario content, the non-conceptual content he attributes to perception, is delimited in this way: it is a way of filling the space around the perceiver, and it is correct when the space around her is filled that way (Peacocke, 1992, Chapter 3). $\{X\}$ is delimited in another space. The boundary of its basin of attraction runs through what the system takes in, not through the ways the world could be.

Error shows the difference. A plastic tomato completes $\{TOMATO\}$, and the expectation fails when it is bitten. Nothing was false. Falsity appears with "This is a tomato." So $\{X\}$ is available before the concept and has no correctness conditions. Content is how $\{X\}$ is described by someone in words.

This places $\{X\}$ with what Hutto and Myin call basic minds without content (Hutto & Myin, 2013).

7.2 Object-word and abstract-word

In a human, $\{X\}$ comes to contain a word: the object-word x^o, learned in the language game first as a motor pattern (the way infants learn to say words or we do when learning a foreign language). The child's $\{TOMATO\}$ contains $tomato^o$.

The object-word is the word used of something present. "This is a tomato," said of the tomato in front of her, uses $tomato^o$. The constellation $\{X\}$ stays pre-verbal. x^o is one more pattern in it and says nothing by itself. The object-word infers nothing, it is simply applied, and it takes part in the expectations of $\{X\}$, as the smell of the tomato leaves does. Animals have the same $\{X\}$ without object-words, this is a common element for basic cognition.

Immersed in the language game, the child acquires a second mode of the same word: the abstract-word x^a, used of things in their absence. "A tomato is red," said with no tomato present, uses $tomato^a$. [14] Most of our language use involves the abstract mode, and this is why it seems difficult at first to grasp what the object-word is about.

$tomato^o$ and $tomato^a$ are one word used in two ways, not two different words. The abstract-word is detached from the occasions in which the word was learned, and it can be written down and read anywhere. The object-word applies; the abstract-word infers.

[14] In earlier work this mode was written x^c and called the concept-word, and property-words were written p^c (Author, 2026). The old name presented the inferential use as the concept, which is the identification this paper rejects, and it clashed with Frege's Begriffswort (section 5). "Abstract" means only "used of no present thing," detached from the occasions and the person in which the word was learned. It is not opposed to "concrete."

7.3 Linguistic-pattern-constellations

The child who says "a tomato is red" with no tomato present has begun to use "tomato" in sentences that do not depend on what is in front of her. Over time she uses it in many such sentences: "Tomatoes are red," "Tomatoes are sweet," "Have you bought tomatoes?" "Tomatoes need sun." Each relates "tomato" to other words. Together they give the word its place among words: the words it goes with, what it follows from, what it excludes, what saying it commits her to. The uses are learned together, and they are completed from a part: hearing "tomato" in a sentence, a speaker is ready for "red" and not for "drown."

The uses of "tomato" that a given speaker makes and would make, in its abstract mode, make up the word's linguistic pattern-constellation, written $\langle tomato \rangle$, or in general $\langle x \rangle$, in lowercase letters.

$\langle x \rangle$ is the counterpart of $\{X\}$ on the side of language. $\{X\}$ is formed from sensory, motor and feeling patterns learned together in encounters with a thing. $\langle x \rangle$ is formed from patterns of words learned together in the language game[15].

One way to see then the abstract-word mode use for "x", is that x^a labels its linguistic pattern-constellation $\langle x \rangle$. In the same way that x^o is a constituent of $\{X\}$, x^a is a constituent of $\langle x \rangle$.

To conclude, we learn words as simple "labels" for object encounters, and at that stage they are part of a pattern-constellation that also includes perception, feeling and motor reaction patterns. Immersed in a language game played by adults we learn to use words in their abstract modes, where the rules learned in training create linguistic pattern-constellation. The same word can still be used in its two different modes; we can see this when at the instant we see a tomato, the word may simply pop-up in our mind and we may say "tomato", with no meaning whatsoever involved.

Inferential roles for words are exclusively stated in x^a mode, and inferences are exclusively belonging to $\langle x \rangle$; Simonelli's inferential role of a word is defined by its linguistic pattern-constellation $\langle x \rangle$.[16]

7.4 Properties

Properties are a specific case of abstract-word mode, because they have only an abstract-word mode and no object-word mode. We write them p^a: there is red^a and no red^o.

While entirely abstract, we learn property-words immersed in a practice, where we learn that what is common in perception for different objects (tomato, sunset, fireman's car, traffic light, etc.) is named using the word "red".

There are theoretical properties for which there is literally no relation to experienced objects, and which are learned exclusively in an inferential manner, inside a language game (usually a specific scientific language game). For example, "entropy" is such a word, that makes full sense only inside a rigorous theoretical framework; we'll get back to this because it is instructive to our case.

[15] There are different functional aspects of how words link to other words in language, which are in fact different sorts of patterns, and not a simple pattern, accounting for the idea of "constellations" used here. But this is a different subject that will not be discussed here.

[16] Inferences are a variety of rule-following, and it is training in Wittgenstein sense that instaurates the rules that govern associations in a specific linguistic pattern-constellation.

7.5 Meaning and public meaning

Each speaker holds a version of linguistic pattern-constellations $\langle x \rangle$, learned from particular people. The public use, which no speaker holds, we note $\overline{\langle x \rangle}$. The public use of "tomato" is $\overline{\langle tomato \rangle}$. Correction runs against the public use. "That is pink, not red" corrects my $\langle red \rangle$[17] against $\overline{\langle red \rangle}$, not against a better speaker's version, even if the correction is made by a certain speaker who does not hold $\overline{\langle red \rangle}$. This can be seen if we ask "why" to the person correcting us, and the answer will be something like "this is the way it is"; this is meant to say that public use terminates into public practice, not in a speaker's definition.

7.6 The linking relation $\mathcal{R}$ relates two uses of words in a claim

We introduce $\mathcal{R}$, named "linking", which coordinates uses of words in their modes. Some examples are:

- General predication: $\mathcal{R}(x^a, p^a)$, "cars are big."
- Singular predication: $\mathcal{R}(x^o, p^a)$, "this car is red."

In a speaker, linking is a learned association. Its normativity comes from correction against $\overline{\langle x \rangle}$, and not from some sort of theoretical rules inside the framework.

A property-word's link is unsaturated, $\mathcal{R}(_, red^a)$, and predication saturates it: $\mathcal{R}(car^o, red^a)$.

7.7 A concept is the link of a word with itself across its two modes, $\mathcal{R}(x^o, x^a)$.

Because here we deal with two different uses for *the same word*, $\mathcal{R}(x^o, x^a)$ is read as "x". For example $\mathcal{R}(tomato^o, tomato^a)$ is read as "tomato", but points to something more complex.

The link itself belongs to neither relatum: $tomato^o$ is in $\{TOMATO\}$, and $tomato^a$ is in $\langle tomato \rangle$, and what $\mathcal{R}(tomato^o, tomato^a)$ means is:

"This tomato (the object I experience as $\{TOMATO\}$) is a tomato (the word I use in certain ways, $\langle tomato \rangle$)"

Neither of its parts can function as concept, because $tomato^o$ acts as a label devoid of meaning, and $tomato^a$ can only be part of the inferential role that we established is what meanings are. I know that at first, saying that the concept tomato is $\mathcal{R}(tomato^o, tomato^a)$ seems difficult to follow, but it makes more and more sense once we see that a concept while linguistic should live on the experience side.

There is an interesting analogy in Russell's definite descriptions (Russell, 1905), where "The present King of France is bald" is decomposed in 3 distinct claims,

1. There is at least one present King of France.
2. There is at most one present King of France.

[17] If a concept is the link between a meaning and experience, a meaning cannot be corrected without correcting the concept as well; the correction of the meaning is explicit, while the correction of the concept is implicit. For example, if I say "strawberries are blue" because I have seen a strawberry painted blue, and I am corrected with "strawberries are red or green, never blue", both the meaning of "strawberry" and its concept are corrected. This can be analysed in more detail, but it is not the point of this paper.

3. Whoever is present King of France is bald.

and becomes

$$\exists x \left(K(x) \wedge \forall y \left(K(y) \rightarrow y = x\right) \wedge B(x)\right)$$

In our case, a concept means saying two things using the same word, and unite them:

1. There is something I experience as tomato, $tomato^o$ as part of $\{TOMATO\}$
2. There is a way I properly use the word "tomato", $tomato^a$ as part of $\langle tomato \rangle$
3. The two modes above, together, are what I call concept tomato

7.8 Property-words and their concepts

A property-word has no object-word, so it cannot be linked with itself across two modes. There is no red^o and no $\{RED\}$: nothing is red without being some red thing. The concept of a property is carried instead by the blank of the unsaturated link $\mathcal{R}(_, red^a)$.

The blank can be filled in two ways:

- Filled with abstract-words, as in $\mathcal{R}(tomato^a, red^a)$, "tomatoes are red," it gives general predications. These are inside $\langle red \rangle$, and they belong to the meaning of "red."
- Filled with object-words, as in $\mathcal{R}(x_i^o, red^a)$, with i running over the things I have seen as red, it gives applications: "this tomato is red," "this car is red."

The concept red is the second set of fillings. What holds them together is the color pattern that the constellations $\{X_i\}$ share. The ripe tomato, the sunset and the fireman's car are different constellations, and red is what is common in them, as described in the section on properties. The concept of a property rides on the constellations of many things.

7.9 Mary has the meaning of "red" and lacks the concept

In the garden, Mary sees the tomatoes. Her color vision works, and her $\{TOMATO\}$ contains the color patterns of tomatoes of every ripeness. She also has general predications for "red": ripe tomatoes are red, red is a color, nothing is red and green all over. What she does not have is any filling $\mathcal{R}(x_i^o, red^a)$. Nothing she has seen was ever tied to the word. So when she is asked to pick the red tomatoes, she cannot do it.

A further inference would not help, because inferences run between abstract-words, and she has those. Her sentience is not missing either, because she sees the tomatoes and they have a different color. What she lacks is the link between what she sees and the word: a filling of the blank with an object-word.

At a closer look, apparently inside the room both Marys have also the concept red $\mathcal{R}(x_i^o, red^a)$. Looking at instances of objects she perceives as gray while someone is saying "this is red", "that is red" it is creating some sort of $\mathcal{R}(x_i^o, red^a)$. But there are two problems with that, one is that the game is rigged and she learns "false" concepts for colors. Once outside, she can see with her eyes that the sensations are different from grays and, moreover, she cannot apply her "concepts" of red and green. The second is that even inside, as we have shown before, she will arrive to say about identical grays that "this is green and that is blue" and cannot pick green from blue because

identical. This second point shows that in fact what she develops in the room are rather pseudo-concepts that she can't apply, precisely because their fundamental nature is inferential and not experiential. What looks like concepts are in fact meanings inffered from language use.

This also changes where the remainder is placed. For Simonelli, Mary lacks no conceptual content: what she lacks is the experience, which belongs to sentience, and anything more would have to be ineffable. In the framework what she lacks is a link, and a link shows in public. It is used each time she says "this is red" of something, and each such application can be corrected by others. So the remainder is not ineffable, and it does not belong to sentience. It belongs to the practice in which words are applied to things.

7.10 The language model has meanings without concepts

The model has no $\{X\}$ for anything. So it has no object-words, and no link can connect its abstract-words to anything it has encountered. Image tokens do not change this. An object-word is a constituent of a constellation formed in encounters that mattered to someone, and an image token is a symbol among other symbols (section 3).

On the side of meaning, the model is better placed than any of us. Its weights are traces of $\overline{\langle x \rangle}$, the public use, which no speaker holds. For "red" it has more abstract-word fillings than any speaker has. This is Simonelli's verdict (p. 5, n. 4), and the framework agrees with it as a verdict about meanings. For perceptual words like "red" and "tomato," the model has the meanings and lacks the concepts, as Mary does.

The meanings the model has were instituted by people who had the concepts. The public use of "red" was formed by speakers applying the word to red things and correcting each other's applications (see "Meaning and public meaning"). For the individual speaker, the concept depends on the meaning, because one relatum of the link is the abstract-word with its place in $\langle x \rangle$. For the community, the meaning depends on concepts. The model inherits the meanings without the applications that formed them.

7.11 Theoretical words: where concept and meaning coincide

For theoretical words, the blank admits only abstract-word fillings, for anyone. $\mathcal{R}(_, entropy^a)$ is filled by $\mathcal{R}(gas^a, entropy^a)$ and the like, and nothing anyone experiences is an object-word for entropy. The physicist's object-words attach to the readings of instruments. Where the blank admits no object-word for anyone, the abstract-word fillings are all the link there is. The link never leaves $\langle x \rangle$, and the concept and the meaning coincide.

For these words Simonelli's semantics is exact. For "entropy," having every inference is having the concept, and the model's "entropy" is as good as the physicist's, or better. Brandom makes a similar point: he holds that the narrow restriction could reconstruct the contents of actual concepts only in some regions of mathematics (Brandom, 1994, pp. 131–132).

"Red" is an abstract-word, red^a, but it is not theoretical, because its blank admits object-words. Mary holds "red" the way everyone holds "entropy." Hyper-inferentialism

describes every word in this way. Its semantics admits only abstract-word fillings, which is exact for theoretical words and leaves out what sighted speakers have for "red."

7.12 The circularity

We can now return to Simonelli's reason for refusing language-entry rules. To say which perceptual state licenses "this is red," we would have to describe it as seeing something as red, and so use the meaning the theory is supposed to explain. The S-rule escapes the circle by putting a sentence about seeing in the place of seeing.

In the framework the perceptual side is described without "red." It is a set of constellations $\{X_i\}$, formed in encounters, with their sensory, motor and feeling patterns, and these have no content. They are identified by how they are formed, in encounters, and not by what they would be about, so describing them needs other concepts, as any description does because it uses words, but not the concept red. Because they say nothing, they cannot justify anything either, so they are not a Given in Sellars's sense. The concept red is neither the perceptual state nor the inferential role. It is the link between them, the set of fillings $\mathcal{R}(x_i^o, red^a)$. The circle does not arise, because nothing on the perceptual side is described through the concept.

This leaves Simonelli's semantics in place. It is a correct theory of meaning, and the S-rule is an elegant part of it. What it cannot register is the link, because a semantics made of sentences has no place for a tie between a word and a constellation.

8 A Sellarsian objection

Sellars could reply that the account of section 7 adds nothing to his. In his account perception is already connected with inference. A trained speaker who sees a tomato responds with "this is a tomato," and this sentence then enters intralinguistic moves like any other, as in "this is a tomato, so it needs sun." The entry transition and the moves are both parts of the functional role of the word, which his dot-quotes classify as •tomato• (Sellars, 1974). On this reply, the two modes of section 7 are the entry transitions and the intralinguistic moves, and the link between them is the role itself.

Part of this reply is correct, and section 7 keeps it. Neither mode is a concept on its own. The object-word alone is a label devoid of meaning, like the trained response "this is green," which Sellars, in §19 of *Empiricism and the Philosophy of Mind* (Sellars, 1956/1997), says is not enough to have the concept green, or Brandom's parrot. The abstract-word alone has a place in $\langle green \rangle$, and that place is a meaning.

The difference lies in where perception meets the word. In Sellars's account it meets it at a sentence. Perception is the occasion of the response, and the word in "this is a tomato" is the same word that figures in inferences. For Sellars the word has one mode, and perception stays outside it. In the framework the word itself has an additional mode linked to non-verbal cognition. The object-word $tomato^o$ is a constituent of $\{TOMATO\}$, learned together with the colour, shape and smell of tomatoes, at first as a motor pattern, and it is at hand whenever a tomato is. It is not a sentence and it makes no transition. This mode is what Sellars's account lacks, and it is what connects the word with non-verbal cognition.

The concept is the link between the two modes, $\mathcal{R}(tomato^o, tomato^a)$, and we read the link as the word itself: "tomato." Both relata are modes of one word, so the concept stays entirely in the linguistic realm, and psychological nominalism holds. There is no concept tomato without the word "tomato," and animals, which have $\{TOMATO\}$, have no concept of it. The link also belongs to a speaker. Sellars's dot-quotes classify the role of a word in a language, which is public. The link ties a speaker's own constellations to the word.

The link is hard to see because ordinary use hides it. "Tomato" said of the tomato in front of me and "tomato" in "tomatoes need sun" are written the same way and look like one word with one use. As said in section 7.7 and worth repeating here, Russell showed that "the present King of France is bald" hides three claims under the form of a simple subject-predicate sentence (Russell, 1905). In the same way, one written word hides two modes of use and the link between them, which can be spelled out as: "this tomato (the object I experience as $\{TOMATO\}$) is a tomato (the word I use in certain ways, $\langle tomato \rangle$)." This also explains why a concept seems at once linguistic and on the side of experience. Sellars is right that it is linguistic, and it is tied to experience through its object-word relatum, which is part of a constellation.

References

Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard Univ. Press.

Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, *42*, e215. https://doi.org/10.1017/S0140525X19000049

Frege, G. (1997). *The Frege reader* (M. Beaney, Ed.). Blackwell Publishers.

Held, R., Ostrovsky, Y., de Gelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, *14*(5), 551–553. https://doi.org/10.1038/nn.2795

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554–2558. https://doi.org/10.1073/pnas.79.8.2554

Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. MIT press.

Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, *32*(127), 127–136. https://doi.org/10.2307/2960077

Peacocke, C. (1992). *A study of concepts*. MIT Press.

Putnam, H. (1975). The meaning of "meaning." In K. Gunderson (Ed.), *Language, mind, and knowledge* (pp. 131–193). University of Minnesota Press.

Russell, B. (1905). On denoting. *Mind*, *14*(56), 479–493. https://doi.org/10.1093/mind/XIV.4.479

Sellars, W. (1954). Some reflections on language games. *Philosophy of Science*, *21*(3), 204–228. https://doi.org/10.1086/287344

Sellars, W. (1962). Truth and "correspondence." *The Journal of Philosophy*, *59*(2), 29–56. https://doi.org/10.2307/2023294

Sellars, W. (1974). Meaning as functional classification: A perspective on the relation of syntax to semantics. *Synthese*, *27*(3–4), 417–437. https://doi.org/10.1007/BF00484606

Sellars, W. (1997). *Empiricism and the philosophy of mind*. Harvard University Press. (Original work published 1956)

Simonelli, R. (2023). How to be a hyper-inferentialist. *Synthese*, *202*(5), Article 163. https://doi.org/10.1007/s11229-023-04382-1

Simonelli, R. (2026). Sapience without sentience: An inferentialist approach to LLMs. *Asian Journal of Philosophy*, *5*(1), Article 48. https://doi.org/10.1007/s44204-026-00400-4

Wittgenstein, L. (1968). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Basil Blackwell. (Original work published 1953)

Wittgenstein, L. (1972). *On certainty* (G. E. M. Anscombe & G. H. von Wright, Eds.). Harper Torchbooks. (Original work published 1969)

Wittgenstein, L. (1977). *Remarks on colour* (G. E. M. Anscombe, Ed.; L. L. McAlister & M. Schättle, Trans.). University of California Press.

Wittgenstein, L. (2009). *Philosophical investigations* (G. E. M. Anscombe, P. M. S. Hacker, & J. Schulte, Trans.; Rev. 4th ed.). Wiley-Blackwell. (Original work published 1953)