



Inquiry

An Interdisciplinary Journal of Philosophy

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/sinq20>

Semantic burden-shifting and temporal externalism

Jussi Haukioja

To cite this article: Jussi Haukioja (2020): Semantic burden-shifting and temporal externalism, Inquiry, DOI: [10.1080/0020174X.2020.1805704](https://doi.org/10.1080/0020174X.2020.1805704)

To link to this article: <https://doi.org/10.1080/0020174X.2020.1805704>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 06 Sep 2020.



Submit your article to this journal [↗](#)



Article views: 97



View related articles [↗](#)



View Crossmark data [↗](#)

Semantic burden-shifting and temporal externalism

Jussi Haukioja

Department of Philosophy and Religious Studies, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

Temporal externalism is the view that the meanings and extensions of linguistic expressions can be partly determined by contingent linguistic and/or conceptual developments that take place after the time of utterance. In this paper, I first clarify what it would take for temporal externalism to be true, relying on the notion of burden-shifting dispositions. I then go on to argue that existing thought experiments give us reason to expect that temporal externalism can be true of some natural kind terms, and present a new thought experiment suggesting it can also be true of some proper names.

KEYWORDS Temporal externalism; semantic externalism; reference; extension

Introduction

Temporal externalism is the view that the meanings and extensions of linguistic expressions can, at least in some cases, be partly determined by events that have not yet taken place; in particular, the actions, judgments, and so on, of future speakers. It is, then, a version of semantic externalism, but one that has received considerably less attention in the literature than natural kind externalism and social externalism. Moreover, many people who are sympathetic to other forms of semantic externalism appear to be more skeptical towards temporal externalism. In this paper, I will investigate what it would take for the view to be true, relying on the notion of burden-shifting dispositions, and suggest that we have good reason to believe that temporal externalism in fact can be true of some of the linguistic expressions we use.

Temporal externalism was first introduced, and defended, by Henry Jackman (1999), and practically all discussions of the view make use of his *Grant's Zebra* thought experiment.¹ Although I am inclined to think his

CONTACT Jussi Haukioja  jussi.haukioja@ntnu.no

¹See Jackman (1999), Brown (2000), Stoneham (2003), Jackman (2005), Collins (2006).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

thought experiment does work, I am going to use another thought experiment to motivate the view here. This thought experiment, due to Gary Ebbs (2000) was not initially meant to establish temporal externalism – ironically, Ebbs mentions temporal externalism as a possible reaction to it, only to dismiss it as ‘outrageously bizarre’. Nonetheless, I think it can be used to do just that. Ebbs’ and Jackman’s thought experiments are structurally very similar, and both feature natural kind terms.² At the end of my paper, I will develop a structurally similar thought experiment, featuring a proper name.

Ebbs’ thought experiment

The factual and historical background for Ebbs’ thought experiment (2000, 248–249) is, firstly, that platinum is very similar to gold in appearance; moreover, both dissolve in *aqua regia*. Secondly, platinum was not in fact discovered until the mid-eighteenth century. At this point, chemists were able to distinguish it from gold; in the seventeenth-century chemists could have relied on solubility in *aqua regia*, and might well have categorized a sample of platinum as ‘gold’.

Against this background, Ebbs invites us to imagine a Twin Earth, completely indistinguishable from Earth up until 1650, when large deposits of platinum are found in Twin South Africa. Given that platinum dissolves in *aqua regia*, the chemists on Twin Earth conclude that platinum should be called ‘gold’. This usage becomes standard, and platinum is widely called ‘gold’, and used in the same way as gold on Twin Earth. Later on, chemistry develops very much as on Earth, with one exception: when new methods become available, the chemists on Twin Earth conclude that there are two kinds of ‘gold’: one with atomic number 79, one with atomic number 78. The Twin Earth term ‘gold’ denotes, then, a disjunction of two elements, gold and platinum (in our terminology).

Now, imagine two speakers: John Locke on Earth, and Twin John Locke on Twin Earth, in 1650. Suppose both utter: ‘There are huge deposits of gold in those hills’, Locke pointing to hills in South Africa, Twin John Locke pointing to hills in Twin South Africa. Both hills contain large deposits of the element with atomic number 78 (now called ‘platinum’ on Earth). If we take speakers’ own judgments of sameness of extension at face value (as Ebbs argues we should), we should say that John Locke’s utterance is false: the metal in the hills in South Africa is not in the extension of his term

²Jackman uses biological kinds, while Ebbs uses chemical kinds. It is precisely because of this difference that I prefer Ebbs’ thought experiment: chemical kinds are far less messy when it comes to the preconditions for introducing a term, and the kinds of judgments concerning sameness of extension that feature centrally in the thought experiments.

'gold'. By the same token, we should say that Twin John Locke's utterance is true: the metal in the hills in Twin South Africa *is* in the extension of *his* term 'gold'. But, we are assuming that Earth and Twin Earth are identical up to 1650: the difference in the extensions of the terms used by the two John Lockes is due to events taking place long after their utterances.

The structure of the argument is quite simple, but worth spelling out in detail, as we will return to it later: in 2019, 'gold' has a different extension on Earth and Twin Earth: on Twin Earth, it applies to the element with atomic number 78, on Earth, it does not. When we look at Earth and Twin Earth in isolation, we, or the speakers on the relevant planets, do not feel any pressure to say that a reference change has occurred at any point: according to our ordinary judgments of sameness of extension over time, the extensions have been stable between 1650 and 2019. Moreover, as the situation is symmetrical, it would be quite problematic to try to argue that reference change has occurred on one of the planets, but not the other, discarding the ordinary judgments of the speakers in one of the communities altogether. The only way to make sense of this is to say that the extensions of 'gold' on the two planets were different already when the two John Lockes used them, in 1650.³ But since Earth and Twin Earth are identical up to 1650, temporal externalism follows.

Internalisms and externalisms

The thought experiment described above seems compelling to me. However, to evaluate the status of temporal externalism more generally, I think it will be helpful to take a step back and ask what, in general, makes specific versions of semantic externalism true or false. In this section, I will review the distinction between *first-order* internalism and externalism on the one hand, and *metainternalism* and *metaexternalism* on the other, that Daniel Cohnitz and I have drawn in joint work (Cohnitz and Haukioja 2013). In the next section, I will employ this distinction and try to cast light on what makes different versions of (first-order) externalism true, if and when they in fact are true.

The familiar opposition between semantic internalism and semantic externalism concerns the question: what determines the meaning or

³Ebbs takes the thought experiment to show that the use of a term does not determine its extension. His notion of 'use', however, only includes the linguistic dispositions and mental states of speakers up until the time of utterance: on a more relaxed notion of 'use' which includes future usage, mental states, etc., we can accept the conclusion of Ebbs' thought experiment without rejecting the *use determines extension* principle (see Collins 2006 for discussion).

extension of linguistic expressions? We can call this the *first-order* question, and the internalist and externalist answers to the question *first-order* internalism and externalism. ‘Determination’ can here be read as a supervenience claim: semantic internalists claim that the meaning and/or extension of linguistic expressions supervenes on the internal properties of the speaker using an expression at the time of utterance, while semantic externalists deny this. Temporal externalism is, then, a first-order externalist view: John Locke and Twin John Locke, in Ebbs’ thought experiment, are internal duplicates in 1650, but according to temporal externalism, their tokens of ‘gold’ are not identical in meaning and extension.⁴

Another question, often not explicitly distinguished from the first-order question, is this: what makes it the case that a particular internalist or externalist answer to the first-order question is correct or not? We can call this the *metaquestion*. Here, too, internalist and externalist views are possible: Cohnitz and I call these *meta-internalism* and *meta-externalism*. These are, then, views about what determines the correct theory of reference for a term. A meta-internalist holds that the correct theory is determined by factors internal to the speaker at the time of utterance, while a meta-externalist will deny this, claiming that the correct theory can at least partly be determined by factors external to the speaker.

In Cohnitz and Haukioja (2013), we argue against meta-externalism, claiming that it fails to make sense of the role that reference, and theories of reference, play in the explanation of successful communication: if meta-externalism were true, there could be systematic mismatches between what linguistic expressions refer to, and what information is successfully communicated by competent speakers using those expressions. We adopt meta-internalism, but at the same time accept first-order externalism: once the first-order question and the metaquestion have been properly distinguished from each other, we can see that the answers we give to the two are logically independent, and there is no contradiction in giving an internalist answer to one while giving an externalist answer to the other. On our combination of views, the reference of at least some terms we use is partly determined by external factors (first-order externalism), but the fact that reference *is* so determined, is determined by factors internal to us (meta-internalism).

⁴There are, obviously, major issues that would need to be clarified here, such as: which features of meaning are claimed to be dependent on external features, how ‘internal’ and ‘external’ are to be understood, and so on. This not the place to go into detail about these issues, for discussion, see Haukioja (2017).

Meta-internalism, as such, does not specify *which* internal features of a speaker determine how her terms refer. Our preferred view, *dispositionalist* meta-internalism claims that the fact whether a given semantic internalist or externalist view is true of an expression, as used by a speaker, is supervenient on that speaker's dispositions to apply and interpret that expression – including the speaker's dispositions to apply and interpret the relevant expressions in conditions that are non-actual, as well as the speaker's dispositions to re-evaluate, retract and revise her usage in response to empirical information. Systematic patterns in such dispositions determine whether a term's reference is based on descriptions, causal chains, and so on: theories of reference and extension are, on this view, theories about patterns of dispositions. A classical descriptivist theory of reference for proper names, for example, claims (roughly) that a speaker's dispositions to apply and interpret a name are only sensitive to the information about what satisfies the descriptions associated to the name by the speaker, while a causal-historical theory will claim (again, roughly), that the speaker's dispositions to apply and interpret the name are sensitive to information about the actual causal origin of the name in the speaker's speech community.

Meta-internalism, in our view, makes sense of why internalism may initially seem attractive: in one sense, meanings and semantic facts are not 'out there' in the world, but due to the way we in fact use and are disposed to use our terms. Yet, in another sense, meanings *are* out there: meta-internalism can be combined with first-order externalism. At least for some terms, we are systematically disposed to let external factors affect the way we use and interpret the terms: we are disposed to 'shift part of the semantic burden' onto external factors, making the meanings of the relevant terms at least partly dependent on those external factors. But, given meta-internalism, the fact that some meanings are partly determined by external factors is not a brute metaphysical fact.

What makes externalism true?

To illustrate how dispositionalist meta-internalism and first-order externalism can be combined, let us first take a look at how more familiar versions of semantic externalism arise and are made sense of within dispositionalist meta-internalism; we can then use the picture that emerges to evaluate temporal externalism. A key notion here is that of *burden-shifting dispositions*, hinted at above. These are second-order dispositions to re-evaluate and retract one's applications of a word, in response to new information. The different versions of first-order externalism are dependent on different

kinds of burden-shifting dispositions: that is, dispositions that are sensitive to different kinds of external factors.

First, take natural kind externalism. Natural kind externalism, if true, is made true by (roughly) speaker dispositions to re-evaluate their application of a term in response to empirical information about the causally relevant underlying properties of samples potentially in the extension of the term. In the case of 'water', for example, we have dispositions to evaluate the correctness of actual and counterfactual applications of the term according to whether or not the term is applied to samples which share the underlying structure of the substance that is causally connected in the appropriate way to our actual usage of 'water' (or something similar: the details will depend on one's preferred theory of reference). We also have dispositions to re-evaluate our application of such terms in the face of new empirical information about what the world is like: were we to find out that most or all of the stuff we have called 'water' in the past has a different underlying structure than we thought, we would adjust our categorization judgments accordingly. This, I take it, is what sets natural kind terms apart from most other general terms: we simply do not have analogous burden-shifting dispositions with respect to terms such as 'bachelor' or 'table'.⁵

Note that these are precisely the sorts of dispositions that standard externalist thought experiments are designed to point our attention to. Sometimes the reliance on burden-shifting dispositions is explicit: a very good example of this can already be found in Putnam (1975, 237–238), where he defends the view that 'gold' (or χρυσός) meant gold, as used by Archimedes, by relying on Archimedes' presumed dispositions to re-evaluate his usage, had he witnessed later scientific experiments and/or been informed about modern theories of molecular structure.

Dispositionalist meta-internalism can, in a very natural way, make sense of why such thought experiments have been found so convincing: the thought experiments direct our attention at the very dispositions which make externalist views true, and which set our usage of the relevant terms apart from that of most other terms we use. If a speaker turned out not to have these second-order dispositions, we would not say that she is using the relevant term with the same meaning as we are: for example, if a speaker consistently rejected the Putnamean judgments about 'water' in response to Twin Earth thought experiments, insisting that XYZ is water, and moreover showed no signs of deferring to expert

⁵Or 'pencil' – here I part ways with Putnam (see Cohnitz and Haukioja 2020 for discussion).

usage, we should, I think, conclude that the speaker is not using ‘water’ with the same externally determined meaning as Putnam and (I take it) most other speakers.⁶

Similarly for social externalism. One familiar variant of social externalism, if true, is made true by speaker dispositions to re-evaluate their application of a term in response to new information about experts’ use of the term. For example, I can refer to elm trees with my term ‘elm’, even though I cannot tell them apart from beech trees, on the basis of my dispositions to defer to people who can (e.g. botanists or gardeners). Should I find that my classification of trees into elms and beeches differs from that of an expert, I would immediately be disposed to revise my earlier application of the terms and align my usage with the expert. Other variants of social externalism can be made true by different kinds of burden-shifting dispositions: for example, I might not be disposed to defer to any particular speakers and their usage, but rather be disposed to align my usage with that of the majority, and further to revise my earlier applications, should I find out that I my usage was not in line with the majority. According to dispositionalist meta-internalism, then, the meanings of the terms I use may depend on how other speakers use them, but only if, and because, I am disposed to defer to other speakers.

Is temporal externalism true?

Finally, we can extend the general idea introduced above to temporal externalism. If we accept dispositionalist meta-internalism, temporal externalism could arise in the following way. Temporal externalism would be true if, for some speaker *S* and word *w*, *S* is disposed to re-evaluate her use of *w*, in response to information about future use of *w* in his/her speech community. Roughly, this means that *S* is disposed to accept and go along with a range of different interpretations of *w*, and retract or not retract accordingly. Of course, in most cases speakers do not in fact receive the relevant kind of information about future use. What matters is the presence of such dispositions, not that they in fact be manifested. In some cases they *can* be manifested, when the relevant events take place within in the speaker’s life span, but in most cases not. However, this is not any more problematic in the case of temporal externalism than with other externalist views: we saw earlier how Putnam relied

⁶Such speakers may well actually exist. As noted above, according to dispositionalist meta-internalism, theories of reference are theories about systematic patterns of dispositions among speakers, and the precise nature of such dispositions is obviously an empirical matter.

on Archimedes' presumed dispositions to react to events that took place more than two thousand years after his death, in arguing for natural kind externalism. (In Putnam's example the future events did not play a role in determining the meaning of 'gold', but merely would have provided Archimedes with information that was not available to him at the time, but this is immaterial: in both cases we rely on speaker dispositions to react to events which, due to practical factors, are all but guaranteed not to happen within the relevant speaker's life span.)

Somewhat more concretely, one way⁷ for temporal externalism to be true would be the following: a speaker or a group of speakers have enough familiarity with an entity, property, or a kind, to introduce a name for it. However, due to limited information, their semantic intentions are less than fully determinate as to which specific type of entity, property, or kind is in question. The resolution of this matter is left to future speakers, through deferential dispositions: the speakers are disposed to accept and go along with a range of different interpretations of the term they have introduced, depending on how future speakers decide to use the term. The situation described here is, of course, precisely the situation of the speakers on Ebbs' Earth and Twin Earth (including Locke and Twin Locke), prior to 1650: the speakers had enough familiarity with a substance to introduce a term, but their semantic intentions were not fully determinate as to whether their term 'gold' should denote a unified kind, or a disjunctive kind.

But, it might be objected, is it really plausible to suppose that ordinary speakers, in 1650, had such sophisticated dispositions to re-evaluate the content of their utterance in response to various possible future events, which they could not even imagine at the time? After all, John Locke and Twin John Locke did not even have the concepts of element and atomic number. I have two responses to this objection. Firstly, speakers do not need to be able to represent the various possible future courses of events, at the time of utterance. What matters is their dispositions to react to such events, were they to take place. Again, the situation is quite analogous to Putnam's reliance on Archimedes' presumed dispositions. Secondly, we are primarily interested in how *our* terms refer, and

⁷I am not claiming that this is the *only* way in which temporal externalism could arise, merely that this is *one* way, and moreover, that this is how it conceivably could arise in a language used by creatures like us. Just as in the case of natural kind externalism and social externalism, there are many different variations of externalist views. I suggested that different forms of social externalism, for example, may arise out of burden-shifting dispositions that are sensitive to different aspects of language use in the speaker's community; similarly, different variants of temporal externalism could conceivably arise out of different kinds of 'forward-looking' burden-shifting dispositions.

whether temporal externalism is true of them. To the extent that we find Ebbs' judgments concerning his thought experiment compelling, we apparently *do* have the kinds of burden-shifting dispositions that would make temporal externalism true.

Temporal externalism for proper names

The above explanation requires that the term in question can be introduced and used, while the speakers are ignorant or mistaken about the criteria for belonging in its extension. Accordingly, temporal externalism can arise in this way only for terms the content of which is not descriptive: the prime candidates are, then, natural kind terms and proper names. As is evident from the above, all existing discussions of temporal externalism have focused on natural kind terms. No one has, to my knowledge, presented a thought experiment that would aim to establish temporal externalism for proper names; I will conclude this paper by doing so. The thought experiment has a similar structure as Jackman's example of 'Grant's Zebra' (and the similarities do not end here, as we will shortly see) and Ebbs' thought experiment discussed above.

In the new thought experiment, a group of explorers come ashore in new, unexplored territory, at time t , and introduce a name for it. Let us suppose that the explorers are led by a man called Grant, and they call the newly found land 'Grant's Land'. They intend the new name to denote this unexplored land, without a definite intention that it should denote (say) a whole island rather than a region of an island, or vice versa: they do not know whether the land area they have found is an island, or maybe connected to other, already known areas of the world.

Suppose the landmass is in fact an uninhabited island, and consider two alternative future courses of events:

- 1: Shortly after t , another group of explorers come ashore in another part of the island and name it 'Chapman's Land'. Eventually, the two groups meet halfway through the island, and the two names are thereafter commonly used to denote the two halves of the island.
- 2: The first group of explorers, led by Grant, charts the whole island before other explorers find it. The name 'Grant's Land' is thereafter commonly used to denote the whole island.

If we consider either of the two alternative courses of events in isolation, ignoring for the moment the possibility of the other course of events, there is, I would say, no pressure to say that reference change has occurred

between the original baptism and later uses. This suggests that in alternative 1, the first uses of the name did not denote the whole island, but rather only a half of it, while in alternative 2, the first uses of the name denoted the whole island. But we are assuming that the two alternatives are identical until shortly after the introduction of 'Grant's Land': the referent of the name is therefore determined by events that had not yet taken place at time t . If so, temporal externalism holds for 'Grant's Land'.

Here's a possible objection⁸: should we not rather say that reference in the original situation gets fixed to the most *salient* or most *natural* uncharted geographical area, in both alternative 1 and alternative 2? If so, the first uses of the name denote the whole island in both alternatives, and the reference of 'Grant's Land' undergoes a change in alternative 1, but not alternative 2.

I have two responses to the objection. I do agree that a term *can* be introduced such that its reference *would* be fixed to the most salient or most natural uncharted geographical area. But, based on the meta-internalist explanation of first-order externalism presented above, this would need to be grounded in the speakers' burden-shifting dispositions: in this case their dispositions to reject and revise uses of the name that do *not* aim to denote the whole island. If we assume that the first users of 'Grant's Land' had such dispositions, the name would indeed undergo a reference change in alternative 1. But we are free to imagine that the explorers did *not* have such dispositions. Given that we can imagine this – as I take it we clearly can – the objection can be disarmed.

Secondly, the thought experiment can be tweaked to disarm the objection. We can, for example, imagine that the land mass is peanut-shaped, and that the two halves of the land mass are connected by an isthmus that is above the surface of the water at low tide, but not at high tide: the details are to be tweaked in whichever way necessary to make the two candidate referents – the whole peanut-shaped landmass, and one of its two halves – equally natural or salient. Now we can rerun the thought experiment, and any inclination to say that the first uses of 'Grant's Land' in alternative 1 should denote the whole landmass should have disappeared.

Conclusion

In this paper I have given an explanation of how temporal externalism could come to be true for some of the terms we use. Moreover, I have

⁸Thank you to Manuel García-Carpintero for raising the objection in discussion.

tried to situate temporal externalism in a more general framework for explaining externalist meta-semantics. I conclude that temporal externalism *can* be true of some of the terms we use, in particular natural kind terms and proper names: we can imagine speakers having the kinds of burden-shifting dispositions that would make their reference dependent on future events in the speaker's speech community. This does not prove that temporal externalism in fact *is* true of some of the terms we use: on my view, this depends on the kinds of systematic patterns of dispositions that we in fact have, and this is an empirical issue: we cannot simply assume that our armchair judgments about our own dispositions, let alone the dispositions of our linguistic peers, are reliable. But I do find the thought experiments compelling, and think temporal externalism *probably* is true of some of our own terms, and hope that the clarifications I have made along the way will contribute to eventually putting the view to empirical test.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- Brown, Jessica. 2000. "Against Temporal Externalism." *Analysis* 60: 178–188.
- Cohnitz, Daniel, and Jussi Haukioja. 2013. "Meta-Externalism vs Meta-Internalism in the Study of Reference." *Australasian Journal of Philosophy* 91: 475–500.
- Cohnitz, Daniel, and Jussi Haukioja. 2020. "Variation in Natural Kind Concepts." In *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*, edited by Å. Wikforss and T. Marques, 128–146. Oxford: Oxford University Press.
- Collins, John. 2006. "Temporal Externalism, Natural Kind Terms, and Scientifically Ignorant Communities." *Philosophical Papers* 35: 55–68.
- Ebbs, Gary. 2000. "The Very Idea of Sameness of Extension over Time." *American Philosophical Quarterly* 37: 245–268.
- Haukioja, Jussi. 2017. "Internalism and Externalism." In *A Companion to the Philosophy of Language*. 2nd and extended ed., edited by B. Hale, C. Wright, and A. Miller, 865–880. Chichester: Wiley-Blackwell.
- Jackman, Henry. 1999. "We Live Forwards but Understand Backwards: Linguistic Practices and Future Behaviour." *Pacific Philosophical Quarterly* 80: 157–177.
- Jackman, Henry. 2005. "Temporal Externalism, Deference, and Our Ordinary Linguistic Practice." *Pacific Philosophical Quarterly* 86: 365–380.
- Putnam, Hilary. 1975. "The Meaning of 'Meaning'." In *Mind, Language and Reality: Philosophical Papers, Volume 2*, 215–271. Cambridge: Cambridge University Press.
- Stoneham, Tom. 2003. "Temporal Externalism." *Philosophical Papers* 32: 97–107.