
BELGRADE PHILOSOPHICAL ANNUAL 2022

FILOZOFSKI GODIŠNJAK



| 35/2

**BELGRADE PHILOSOPHICAL ANNUAL/
FILOZOFSKI GODIŠNJAK 35(2)/2022**
Institute of Philosophy, Faculty of Philosophy,
University of Belgrade
Belgrade, Čika Ljubina 18–20

Belgrade

Year XXXV

YU ISSN 0353-3891

UDK-1

Editor

Voin Milevski (University of Belgrade)

Associate Editors

Miloš Arsenijević (University of Belgrade)

Jovan Babić (University of Belgrade)

Leon Kojen (University of Belgrade)

Živan Lazović (University of Belgrade)

Timothy Williamson (University of Oxford)

Slobodan Perović (University of Belgrade)

Editorial Board

Berit Brogaard (University of Missouri, St. Louis)

Paul Boghossian (NY University)

Aleksandar Jokić (Portland State University)

Jan Narveson (University of Waterloo)

Georg Meggle (University of Leipzig)

J. Angelo Corlett (San Diego State University)

Howard Robinson (Central European University)

Managing Editor

Petar Nurkić (University of Belgrade)

petar.nurkic@f.bg.ac.rs

Belgrade Philosophical Annual is published twice a year and is available online at
<http://www.f.bg.ac.rs/bpa/index.html>

Printed by

Službeni glasnik, Belgrade

This issue is financially supported by

Ministry of Education, Science and Technological Development of the
Republic of Serbia

The statement on publication ethics can be found at the journal website
<http://www.f.bg.ac.rs/bpa>

BELGRADE PHILOSOPHICAL ANNUAL 35(2)/2022

ARE THERE ESSENTIAL INDEXICALS?

Guest editor: Vojislav Božičković, University of Belgrade

John Perry	Are There Essential Indexicals?.....	7
Miguel Ángel Sebastián	Conscious Perception in favour of Essential Indexicality.....	13
Andrija Jurić	The Essential Indexical and Self-Consciousness: 'I', 'Now', and 'Here' as Aspects of Self-Consciousness	31
Peter Ludlow	Cognitive Dynamics: Red Queen Semantics versus the Story of <i>o</i>	53
Andreas Stokke	<i>De Se</i> Thinking and Modes of Presentation.....	69
Eduarda Calado Barbosa	For Their Eyes Only.....	89
Ernesto Perini-Santos		
María de Ponte	Demystifying the myth. Perry: Revisiting the Essential Indexical.....	107

ARE THERE
ESSENTIAL INDEXICALS?



ARE THERE ESSENTIAL INDEXICALS?

Keywords: *Indexicals, Castañeda, Kaplan, character, context, content, self-informative, self-effecting, efficient, eternal.*

(October 25, 2022)

1. I am pleased that the *The Belgrade Philosophical Annual* is devoting an issue to the question, “Are there Essential Indexicals?” In this paper I explain what I meant by “essential indexicals” in “The Problem of the Essential Indexical,” (1979) and two other essays I wrote about the same time,¹ and claim that, at least in this sense, there are some of them.

2. Hector-Neri Castañeda called ‘I’ and ‘now’ “essential indexicals” because they cannot be defined with other expressions. ‘Here’ doesn’t quite make the list; it’s where *I* am *now*. I borrowed the term for something a bit different. Sometimes when we use an indexical to refer to some object it conveys information about that object that other ways of referring to it would not. The indexical is essential (or at any rate very useful) for conveying that information. Consider Jane. She sits in her office a bit before noon on Wednesday. The Promotion Committee, of which she is a member, has a meeting starting at noon that day, in a room just down the hall. Jane plans to attend the meeting on time. With her is her office-mate Fred. Fred knows that Jane has a meeting that day which she plans to attend, although he doesn’t know when it starts. Neither Fred nor Jane has been paying much attention to the time. About a minute before noon, Fred asks Jane, “When does the Promotion Committee meeting start?” Jane responds,

(1) The Promotion Committee meeting starts at noon.

Then she glances at her watch, and sees, to her surprise, that it is just a few seconds until noon. Then she dashes off, telling Fred,

(2) The Promotion Committee meeting starts now.

Jane’s utterance (1) did not explain her dashing off. Her utterance (2) did. The difference was that she used “now” in (2). ‘Noon’ in (1) and ‘now’ in (2) both

¹ Perry, 1977, 1980.

refer to a certain time, 12 p.m. Wednesday. So (1) and (2) seem to express the same singular proposition, that that time has the property of being when the Promotion Committee meets. Jane's using "now" to refer to 12 p.m. Wednesday seems essential to (2)'s providing Fred with an explanation for her departure. In the terminology of my essay, it is an "essential indexical."

3. Is this puzzling? David Kaplan has given us a semantics and logic for indexicals we can use to see.² I summarize:

- i. *Characters* is David Kaplan's term for the meanings of the sort indexicals have. A *character* is a function from *contexts* to *contents*.
- ii. *Contexts* consist of *agents, locations, times, and circumstances*. Intuitively, these are the speaker, time, location and circumstances of an utterance.³
- iii. The *content* of an expression is the object it refers to.
- iv. The *content* of a sentence containing indexicals is a singular proposition about the referents of the indexicals.
- v. The characters of 'I', 'here', and 'now', respectively are functions from a context to the agent, location and time in it. The content of 'today' is the day of which the time of context is a part.
- vi. The characters of demonstratives like 'this' and 'that' and some indexicals like 'you' and 'we' are trickier, but Kaplan's theory gives us what we need to consider the alternatives

4. Back to (1) and (2).

1. The Promotion Committee meeting starts at noon.
2. The Promotion Committee meeting starts now.

I'll assume that "noon", as used here, is an indexical that refers to 12 p.m. of the day of an utterance of it occurs. So, from (1) Fred can reasonably infer Jane will head off to the meeting some time during the day. But from (2) he can reasonably infer that Jane will rush off at the time of her utterance. So, it seems that 'now' in (2) is an essential indexical, in the sense that if (2) had contained another expression or phrase with the same reference he could not have made that inference.

5. But what about Jane? A number of readers thought I meant that the indexical 'now' was essential to Jane's having the belief that motivated her to run off to the meeting, perhaps a sort of mental indexical, part of a "language of thought". But that's not what I claimed.

2 Kaplan, 1979a, 1979b, 1989.

3 Kaplan does not present his account as a theory of utterances. For the purposes of logic, it's better to have a theory of contexts, contents etc. that can be applied to utterances. A sentence can be true in a given context, even if there is no utterance of that sentence that context.

I did claim that Kaplan's concept of a character can help us understand belief-states and other mental states. When any animal perceives, it gets information about what is happening inside and outside of it at the time and place of perception. Episodes of vision, smell, and hearing provide information about the objects in its vicinity that it sees, smells and hears at the time of the episodes. Episodes of visceroreception and more generally interoception provide information about what's going on in the perceiver's body, and introspection provides information about what is going on in the agent's mind. But there need be nothing like an inner word or "indexical idea" involved. The contents of our most basic sorts of mental states, perceptual states of various sorts, are propositions about what is going on in and around the agent at the time of perceptions; no special indexical words or ideas are needed to mark this.

In general, external perceptions — vision, touch, hearing, smell etc., — are normally ways for an agent to know about other things, while interoception and introspection are ways for an agent to know about itself. But external perception is also self-informative. Sitting at a table in a bar, I see a waiter arrive and place a full mug of beer. I learn about the waiter, the table, and the mug. But I also learn something very important about myself: I have a full mug of beer in front of me. Then I will perform a "self-effecting" action. I'll extend my arm, grab the mug with my hand, bring it to my lips and drink from it and nourish myself, or inebriate myself, or both.

6. Our perceptions and our actions share an important similarity with indexicals, what Jon Barwise and I called "efficiency" (1983). Consider the type of action just described. It's not just a way for me to drink a beer in that particular bar. It's a way for anyone with arms and hands and lips to drink a beer in a circumstances like the one I was in. Similarly, indexicals allow different people at different times and places in similar circumstances to say different things with the same sentences. The structure of human beings and bars allows different people at different times and places to perform the same movements, with the results that different people are nourished by different mugs of beer. We all scratch our different backs in the same way, and relieve different itches. So Kaplan's concept of characters can be generalized from types of utterances to types of actions and thoughts. The action of putting one's arm behind one and rubbing yields a result in a certain context: the agent relieves the itching sensation that the agent has at that time and place. Mother Nature clearly appreciated efficiency; it allowed her to use the same design to create millions of creatures of a given species, each one of which had to know about and act productively in different places at different times. It is hard to imagine evolution, or education, or mass production of any sort getting by without efficient design.

7. Efficient sentences, using indexicals and other context-sensitive devices like tense, can be contrasted with *eternal sentences*, which express the same

propositions whenever they are used, and whoever uses them, and whatever the particular circumstances of the use. Frege thought that a “finished science” would consist of true eternal sentences, and that the principles of logic are clearest when we confine ourselves to them. Perhaps for this reason, the study of efficient language wasn’t really a main topic of interest to philosophers of language for the first part of the last century. Frege’s *Begriffsschrift (Concept Language)*⁴ was intended to be a perfect language. It did not contain indexicals, demonstratives or tense.

8. This all changed with David Kaplan’s work on indexicals and demonstratives. Sentences containing such context-sensitive expressions are typically efficient, as Kaplan’s theory explains and predicts. For me, uttering “I’m from Nebraska” is a way of expressing a true proposition, but for most speakers it expresses a false one — with some stellar exceptions like Marlon Brando, Henry Fonda, Richie Ashburn; Jeff Pelletier and Saul Kripke. Once a year, on December 31, uttering “It’s New Year’s Eve” is a way of saying something true, but not for the rest of the year.

Kaplan’s candidates for the characters of common indexicals were not exactly revolutionary. Once we have his system, it seems pretty obvious that the character of ‘I’ should deliver the speaker as referent, ‘now’ the time, and ‘here’ the place. In less obvious cases, like ‘this’ and ‘that’, his theory shows us where to locate the relevant issues. Is the referent the object the speaker is attending to, or the one to which he intends his audience to attend to, or something else? The revolution was not in the specific candidates for characters, but in the whole idea that languages with indexicals and demonstratives have a semantics and a logic, much less a semantics and logic important enough to intrigue a great logician and give philosophers a lot to think about.

9. I’ve tried to explain what I meant by “essential indexicals”. I think there are lots of essential indexicals in this sense, and don’t quite know why anyone would think otherwise.

But this doesn’t seem what most critics of my essays seem to have taken me to mean. I think the examples with which I began “The Problem of the Essential Indexical” — the messy shopper, the lost hiker and the tardy professor — were somehow so engaging that readers immediately felt they knew what I was getting at and what was wrong with it, without reading.⁵ Well, that’s philosophy, the profession I love. I have written a monograph, *The Essential Indexical Revisited* (2020) which goes into some detail criticizing some of these critics.

4 Frege, 1879, 1967

5 See, Millikan 1990, Cappelen & Dever, 2013.

References

- Almog, J., Perry, J., & Wettstein, H. (1989). *Themes From Kaplan*. New York: Oxford University Press.
- Barwise, J., & Perry, J. (1983). *Situations and Attitudes*. Cambridge, Mass.: Bradford-MIT, 2nd edition (Stanford, CSLI Publications (1999)).
- Cappelen, H., & Dever, J. (2013). *The Myth of the Essential Indexical* Oxford. Oxford University Press.
- Castañeda, H. N. (1989). *Thinking, Language and Experience*. Minneapolis: University of Minnesota Press.
- Castañeda, H. N., Hart, J. G., & Kapitan, T. (eds.) (1999). *The Phenomeno-Logic of the I: Essays on Self-consciousness*. Bloomington: Indiana University Press.
- Frege, G. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle.
- Frege, G. (1882). "Über Sinn und Bedeutung." *Zeitschrift für Philosophische Kritik*, NF 100, pp. 25–30.
- Frege, G. (1960). "Sense and Reference." *Philosophical Review* 57 (3), pp. 209–30. A translation by Max Black.
- Frege, G. (1967). *Begriffsschrift, a Formula Language, Modeled Upon That of Arithmetic, For Pure Thought*. Translation of (Frege, 1879) by Stefan Bauer-Mengelberg, in (van Heijenoort, 2002), pp. 1–82.
- French, P. A., Uehuling, T. E. Jr., & Wettstein, H. K. (1979). *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minnesota Press.
- Kaplan, D. (1979a). "On the logic of demonstratives." *The Journal of Philosophical Logic* 8, pp. 81–98.
- Kaplan, D. (1979b). "Dthat." In French et. al., pp. 383–400.
- Kaplan, D. (1989). "Demonstratives." In Almog et al., pp. 481–563.
- Millikan, R. (1990). "The Myth of the Essential Indexical." *Noûs* 24 (5), pp. 723–734.
- Perry, J. (1977). "Frege on Demonstratives." *Philosophical Review*, 86 (4), pp. 474–97. Reprinted in Perry, J. (2000), Chapter 1.
- Perry, J. (1979). "The Problem of the Essential Indexical." *Noûs* 13 (1), pp. 3–21. Reprinted in Perry, J. (2000), Chapter 2.

- Perry, J. (1980). "Belief and Acceptance." *Midwest Studies in Philosophy* 5, pp. 533–42. Reprinted in Perry, J. (2000), Chapter 3.
- Perry, J. (2000). *The Problem of the Essential Indexical*, Expanded edition. Stanford: CSLI.
- Perry, J. (2020). *Revisiting the Essential Indexical*. CSLI Publications.
- van Heijenoort, J. (2002). *From Frege to Godel: A Source Book in Mathematical Logic, 1879–1931*.



CONSCIOUS PERCEPTION IN FAVOUR OF ESSENTIAL INDEXICALITY

Abstract: *It has been widely acknowledged that indexical thought poses a problem for traditional theories of mental content. However, recent work in philosophy has defied this received view and challenged its defenders not to rely on intuitions but rather to clearly articulate what the problem is supposed to be. For example, in “The Inessential Indexical”, Cappelen and Dever claim that there are no philosophically interesting or important roles played by essential indexical representations. This paper assesses the role of essential indexicality in understanding the content of perception and argues that, if the perceptual experiences of different subjects can have the same phenomenology, and thus share a representational content (phenomenal content)—assumptions shared by a well-established research project, including a wide variety of theories—, then phenomenal content exhibit a particular kind of essential indexicality (perspectival). I also discuss how the argument is immune to Cappelen and Dever’s objections to the idea that the content of perceptual experiences is essentially indexical.*

Keywords: *Conscious perception, perspectival content, de se representation, essential indexical, phenomenal information.*

1. Perception and the Problem of Essential Indexical

It is widely accepted that we sometimes represent the world from a point of view. For example, I believe that I live in Mexico City, desire that it doesn’t rain now, and perceive that there is a cup of coffee in front of me. The content of these representations is typically captured deploying an indexical expression. Indexical expressions are expressions that seem to have different references in different contexts in which the content is expressed. Examples of indexical expressions are ‘I’, ‘me’, ‘mine’, ‘now’, ‘tomorrow’, ‘some time ago’, ‘here’, ‘over there’, ‘this’, ‘that’, etc.¹ We can call ‘indexical thoughts’ to those mental representations whose content is expressed by deploying an indexical expression.

Perry (1979) and Lewis (1979) argued that there are indexical ways of representing the world that cannot be reduced to non-indexical ways

1 In this paper I remain neutral on possible reductive relation between indexicals; e.g. ‘here’ as equivalent to ‘my location’.

of representing. Perry and Lewis focus on beliefs and hence, assuming the canonical way to individuate thoughts² in terms of an attitude (e.g. believe) toward a content, the issue can be expressed in terms of the content.³ Indexical thoughts have been taken to be special because traditional theories of content cannot offer a proper characterization of them. They have a special content that is not adequately expressed in the absence of an indexical expression, and for this reason they are often called ‘Essential Indexical’. However, there is a recent agreement that Perry’s and Lewis’ examples are insufficient to establish that a special kind of content must be attributed to indexical thoughts and that there is such a thing as *essential indexicality*.⁴ Cappelen and Dever (2013; henceforth C&D) summarize the idea when they claim that “the considerations coming out of the Perry-Lewis tradition give us no reason to change our theory of content, and provide no evidence that there are philosophically interesting or important roles played by non-constant characters.” (p. 16) The challenge for those who think that there is essential indexicality is to provide arguments that show that there is some philosophically interesting role that indexical representations play that cannot be played by non-indexical ones, and that such representations are “special” in that classical theories fail to capture the nature of their content.

If we look for the need of indexical representations, it is natural to think of the first-person point of view in our conscious experiences. Moreover, focusing on our conscious experience, and perception in particular, is especially important because, as C&D stress, if there are essential indexical thoughts we should expect an essential indexical element in the representations that work as input of other thoughts. Otherwise, if “all of your perceptual contents were objective and non-perspectival, but [...] the beliefs you formed on the basis of those perceptions were frequently perspectival [...] how are your perceptual states going to justify beliefs which systematically turn out to be about something quite different from what is perceptually presented?” (p. 140). C&D devote chapter 8 to deal with the content of perception. They note that although the perspectival nature of conscious perception is often assumed, there are not many arguments in the literature intending to show that its content is special and to what extent.

2 See Fodor (1987, p. 11). According to such a canonical way to individuate thoughts, my belief and my desire *that it rains now* have the same content and their difference is explained in terms of a difference in attitude, whereas differences among thoughts that involve the same attitude, belief for example, are explained by a difference in content.

3 This would not be Perry’s preferred way to present the discussion because he makes a different use of the technical term ‘content’ abandoning the canonical way to individuate thoughts. Nothing hinges in the preferred use of the terms. For discussion of the relation between indexical thought content and Perry’s *belief states* see Torre (2018).

4 See e.g. Boer and Lycan (1980), Cappelen and Dever (2013), Devitt (2013), Douven (2013), García-Carpintero (2017), Ninan (2016), Magidor (2015), Shaw (2020), Spencer (2007), Torre (2018). See Morgan and Salje (2020) for a recent review.

This paper contributes to show that the content of perceptual experience is, at least some-times, essentially indexical. For this purpose, in section 2 I review C&D's discussion of the arguments in favor of essential indexicality from perception. In section 3, I discuss a refined version of those arguments in terms of the current dispute over essential indexicality and, in section 4, I critically assess some possible replies that the skeptic might offer. I conclude that C&D have misestimated the role that conscious perception plays in the vindication of essential indexicality.

2. Arguments from Perception

If our perceptual experiences convey the presence of mind-independent entities, they do so, at least typically, relative to the subject's location. The content of our perceptual experience is often specified by the use of indexical expressions like 'here', 'over there', 'in front of me', 'behind me', 'now', 'a moment ago', etc.⁵ Following Siegel (2011), C&D agree that we naturally describe perceptual content using indexicals, but they note that the question at stake is whether perceptual content *must* be specified that way. C&D claim that it can be expressed by non-indexical means.

Consider the following situation. John and Berit are in different locations but they hear a qualitatively identical sound coming from their left. Herman is in the same location as Berit but facing her. Herman hears the sound coming from his right. The challenge that C&D take to face is to express what is intuitively common between John and Berit and what is different between Berit and Herman. According to C&D, we do not need indexical expressions to characterize any of this. Accounting for the difference is simple, and there is no problem to attribute different non-indexical contents to Berit's experiences—such-and-such sound is coming from Berit's left—and to Herman's experience—such-and-such sound is coming from Herman's left. C&D think that this sort of content attribution can also explain what is common between Berit and John: "each is an x such that x hears S as coming from x 's right" (p.144). This explanation denies that Berit's and John's experience have a common content and explains what is common instead in terms of a common relation to the contents they have. But they remark that, even if one insists that we should explain what is common in terms of a common content, there is still no argument for essential indexicality.

An attempt to offer such an argument is found in Brogaard's (2009) reply to Cappelen and Hawthorne (2009). Brogaard argues that propositions do not instantiate the property of truth or falsity simpliciter but rather relative truth or falsity. The interesting point for current purposes is her defense

5 Those who think that there is a *sense of mineness* constitutive of phenomenology would be inclined to think that phenomenal content is always first-personal; for recent discussion see Guillot and García-Carpintero (forthcoming); Sebastián (2012, forthcoming).

of the claim that the content of perceptual experience does not instantiate such a property. Brogaard offers two arguments in favor of this claim, the first one resting on the empirical assumption that there is variability in color perception and the theoretical assumption that weak representationalism is true; and the second one resting solely on weak representationalism. I will focus on the later one for it has weaker premises.⁶ According to Brogaard, weak representationalism “is the view that the phenomenology of perceptual experience *determines* the content of perceptual experience” (p.221, my emphasis). A position she takes to be endorsed by author like Chalmers (2004), Siegel (2010) and Tye (2002)—something that C&D dispute (ch.8 fn.4)—and to be “exceedingly plausible” (p.222). She presents the argument in the following paragraph:

When I look at two trees at different distances from me, I can see that one tree is further away from me than the other. Moreover, it is plausible that you and I can have perceptual experiences with the same phenomenology of the two trees (perhaps at different times). By weak representationalism, our experiences have the same content. So, our perceptual experiences cannot contain you or me in the content of perception. Rather, they must contain semantic values that have extensions only relative to perceivers. Hence, the contents of our experiences do not instantiate the fundamental monadic properties of truth simpliciter and falsity simpliciter. (ibid. pp. 221–222)

If the argument is compelling, it remains to be seen what the relation between a content not instantiating the property of truth simpliciter and essential indexicality, because as C&D stress “what is important here is not whether we place the perceiving subject “in the content” or not, but whether that perceiving subject has to be represented in a distinctly *de se* or indexical way. Nothing in the arguments here even starts to motivate the latter idea” (p.148).

To illustrate their point they invite us to consider Maeve, who has a non-standard perceptual system sensitive to how things are around someone else, Brigid, rather than how things stand around her (perhaps via a camera attached to Brigid feeding a signal to Maeve’s brain). C&D argue, on the basis of this case, that there is no need for a distinctively first-person feature in the perceptual content. They affirm that Maeve’s experience represents how things are around Brigid. They take this example to show that mere indexicalized truth is not enough for essential indexicality, because indexed truth can be interpreted non-indexically —Maeve can have perceptual contents that have

⁶ A similar argument is attributed to Chalmers (2006) by C&D. Chalmers remarks, in relation to the *de se* phenomenon, that his argument demands that “the perceiver is picked out under an indexical mode of presentation that can be shared between two different perceivers” (p. 60). However, as the skeptic argues, this is insufficient to establish the need of a special kind of content, and Chalmers himself notes that the argument does not depend on the use of centered worlds semantics to characterize phenomenal content.

no absolute truth conditions, but whose role in Maeve's perceptual system is to be true relative to Brigid.

C&D offer a similar reply to Shroeder and Caplan (ch. 8, fn. 6). Schoder and Caplan (2007) argue that subjects can have experiences that are phenomenologically indistinguishable but with different veridicality conditions. In this case, explaining sameness in phenomenology in terms of sameness of non-indexical content is problematic. In reply, C&D propose a lambda extracted content for explaining sameness in phenomenology. In the previous example, the proposed content would be λx . *there is such-and-such sound to the left of x* , a predicate that denotes the property of *being an x such that there is such-and-such sound to its left*. They deny that such a content is a carrier of veridicality, leaving no reason to introduce indexicalized truth-relative content. Moreover, as they claim in reply to Brogaard, even if indexicalized contents were introduced, it is not clear that they amount to the introduction of anything essentially indexical:

True essential indexicality would, for example, distinguish between the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were true relative to Maeve and the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were truth relative to herself. Only the latter would have the *de se* element characteristic of the (putatively) essentially indexical, and nothing in the various arguments from perception imposes a requirement specifically for the latter. (p.149)

Attending these issues and making explicit the relation between indexicalized truth, the subject being placed in the content, and essential indexicality is precisely the aim of the next section. I will come back to the objections of C&D in section 4.

3. A Refined Version in Light of the Current Debate

3.1 *A characterization of the Current Debate*

As I have remarked, the challenge for those who think that there is essential indexicality is to show that there is some philosophically interesting or important roles played by indexical content that non-indexical content cannot play, and that such indexical content is special in that traditional theories of content cannot accommodate it. To begin with, we should get clear about how “traditional theories” of mental content are characterized. There is plenty of disagreement among traditional theories regarding the kind of entities that contents are supposed to be: Fregean senses, classes of possible worlds, complexes of objects and properties, sentences in the language of thought, etc. However, traditional theories agree that contents are the kind of entities that have truth value and satisfy the following two conditions (Kölbel, 2013, Ninan 2016, Perry, 1979):

Global: The truth value of the content depends only on the way the world turns out to be; i.e., given a complete description of the world, we can assign contents a truth value. In this sense, contents have a *global* or *absolute* truth value, they are assigned the same truth value for any location or point of evaluation (e.g. space, time or subject) within a world.

Portable: Mental representations with the very same content can be entertained by different subjects, at different places and different times. In this sense they are *portable* or *shareable*.

To a first approximation we can think of essentially indexical contents⁷ as those contents expressed by an indexical expression that are either NOT-GLOBAL or NOT-PORTABLE, and hence cannot be accommodated by traditional theories of mental content. Those who are skeptic about essential indexicality (e.g. C&D; Magidor, 2015) maintain that there is no reason to abandon either PORTABLE or GLOBAL. They acknowledge that there is a difference in cognitive significance between my belief that I live in Mexico City and my belief that Sebastián lives in Mexico City (where I happen to be Sebastián). However, they claim that this case is no different from other cases of change in cognitive significance involving co-referential terms, as it happens in Frege's famous example of Hesperus and Phosphorus—cf. Perry 2001b. On the other hand, those who think that there is such a thing as essential indexicality fall into one of two groups depending on whether they reject GLOBAL or PORTABLE. We can call 'globalists' to the former and 'portabilists' to the latter.

The globalist agree with the skeptic that indexical thoughts involve differences in cognitive significance of co-referential terms, but they think that PORTABLE should be abandoned because essential indexicality involves some sort of exclusive access. A classical illustration of this view is Frege's idea that in first-person thought the subject is presented to themselves in "a particular and primitive way" in which they are presented to no-one else (1956, p. 298). For analogous reasons, the globalist view is also illustrated by those who think that indexical content is a reflexive singular content—in the previous example something like the singular proposition expressed by *the thinker of this very thought lives in Mexico City*—because they deny that two different thoughts can have the same reflexive content—e.g. García-Carpintero (2017); Perry (2001a, 2001b). The defense of a globalist view depends then on there being philosophically interesting or important roles played by content that can be entertained exclusively from a particular location in logical space.

Alternatively, the *portabilist* rejects GLOBAL. For example, Lewis thinks that when you and I have the belief that we would express by uttering 'I live in Mexico City', we have beliefs with the same content. The content is then

⁷ For discussion of the relation between indexical thought and content see Torre (2018).

a property—that of living in Mexico City—that we both self-attribute. This understanding of the content of belief in terms of self-attributed properties can be translated into a canonical relation between a subject and a truth evaluable content, as Lewis remarks, simply by letting contents be classes of centered worlds—ordered pairs of worlds and centers.⁸ Now, if you do not live in Mexico City, then I have the property that I am self-attributing but you are not, and your belief is false while mine is true. Considering that we both inhabit the same world, global must be rejected because fixing the way the world turns out to be does not suffice to fix the truth value of the content. It is worth stressing that, in this case, the entity that one would take to be the reference of the indexical expression cannot be part of the portabilist content, because the entity that one would take to be the reference of 'I' in your case and mine would be different; and if it were part of the content we could not share the content. In this sense, one can think of this content as involving a perspective that different particulars can occupy, and we can call indexical content that is PORTABLE and NOT-GLOBAL '*perspectival*'. The defense of a portabilist view—the existence of perspectival content—depends then on there being philosophically interesting or important roles played by this sort of not-referentially anchored content. Lewis offers one such argument based on sameness in belief of physical duplicates but such an internalism is rejected by C&D. In the next section, I discuss an argument in favor of perspectival content based on the content of conscious perception.

3.2 *The Revised Argument from Phenomenology*

Along the lines of previous work (Sebastián 2022, MS), I am going to argue that phenomenal content is, at least sometimes, *perspectival* (PORTABLE and NOT-GLOBAL) on the basis of widely accepted premises regarding the phenomenology of experience and its relation to content, focusing on perceptual experiences whose content is expressed deploying an indexical expression. This requires certain clarification to begin with.

By the *phenomenology* or *phenomenal character* of an experience I refer to *what it is like for the subject* to have the experience (Nagel, 1974/2002). In the intended use of the term, phenomenology is constitutive of experiences in general and perceptual experiences in particular (Siegel, 2011). Two token experiences E_1 and E_2 belong to the same phenomenological type if and only if E_1 and E_2 have the same phenomenology. And two token experiences are considered to have the same phenomenology just in case they are phenomenally indistinguishable in ideal conditions—abstracting away from the discriminatory and memory limitations. Now, I take it to be part of our pre theoretical understanding of phenomenology that two numerically different experiences of two different subject can have the same phenomenology. Sure, the characterization of sameness in phenomenology in terms of

8 See for example Ninan (2012) for a more detailed articulation of centered propositions.

indistinguishability invites questions in the interpersonal case (Frege, 1956; Schlick, 1959; Stalnaker, 2000—for detailed discussion see Shoemaker, 1982, 1996). But although different theories disagree with regard to the conditions that two individuals must satisfy in order to have experiences with the same phenomenology, hardly any theory denies such a possibility—think of the experience of two physical duplicates in physically identical circumstances.

Moreover, it is widely accepted that there is an intimate relation between phenomenology and content. Our conscious experiences are a primary source of information; they convey certain information to us, something acknowledged, as we have seen, by skeptics like C&D. The view that experiences have content is widely accepted, especially in the case of perception—see e.g. Byrne (2009), Nanay (2014), Pautz (2010), Schellenberg (2018), Siegel (2010); cf. Travis 2004. This relation has been exploited to theorize on the phenomenology of experience by investigating the content of experience: what experiences convey to us. The idea behind this weak representationalist research project is that there is a representational content associated with the phenomenological type, because “given a specific phenomenology, it seems that if a mental state has this phenomenology it must also have a certain specific representational content” (Chalmers, 2004). A content that we can call ‘phenomenal content’. Consequently, if two experiences have the same phenomenology then they have the same phenomenal content. Unlike Brogaard characterization of weak representationalism, this entailment is neutral on the existence of any dependence relation between phenomenology and phenomenal character. Hence, it comes as no surprise that it has been defended and endorsed by philosophers with very different theoretical perspectives otherwise disagreeing about substantial issues. It can be accepted by those who think that the phenomenology of experience is not exhausted by phenomenal content, and those who think that experiences can have contents that go beyond phenomenal content. Defenders include those who think that a state having certain content (metaphysically) depends upon its having certain phenomenology, and hence explain phenomenal content by appeal to phenomenology; and also strong representationalist who think that conscious states are representational ones and that a state having phenomenology (metaphysically) depends upon its having certain content.⁹ Note that this kind of strong representationalist would deny that phenomenal content is grounded in—or metaphysically depends on—phenomenology.¹⁰ I take these different research projects that defend and endorse the entailment between phenomenology and phenomenal content

9 E.g. Carruthers (2000); Chalmers (2004); Dretske (1995); Kriegel (2002, 2009); Horgan and Tierson (2002); Rosenthal (2005); Sebastián (2022); Siewert (1998); Tye (1997, 2002).

10 Hence, it is expected that Tye, who has defended such strong representationalism, does not endorse weak representationalism as presented by Brogaard—as C&D claim he did in personal communication.

to be “philosophically interesting” and motivated independently of debates on essential indexicality.¹¹

The argument in favour of phenomenal content being perspectival (PORTABLE and NOT-GLOBAL) can be illustrated with the case presented by C&D that we have discussed in the previous section. In this example, Berit is hearing a sound coming from her left. We have seen that it would be natural for Berit to express the content of her perceptual experience by deploying an indexical expression, for example “there is such-and-such sound coming from my left”. But as we have seen this, *per se*, is silent on the problem of essential indexical as C&D argue. Now, let John be in a different location having a different token experience with the same phenomenology as Berit. He would also express the content of his experience as “there is such-and-such sound coming from *my* left”, which again is silent on whether the content is essentially indexical. However, if sameness in phenomenology entails sameness in phenomenal content and Berit’s and John’s token experiences have the same phenomenology, then these two token experiences must have the same phenomenal content. This rules out a globalist reading of the content because portable has to be respected to make room for the possibility of John and Berit having experiences with the same phenomenology if the entailment between phenomenology and phenomenal content is accepted. The phenomenal content of Berit’s experience cannot be adequately characterized in terms of restricted access for a particular individual, as Frege for example suggests, because both Berit and John must have representations with the same phenomenal content when they have experiences with the same phenomenology. Phenomenal content is not adequately expressed by “there is such-and-such sound coming from Berit’s left” either, because it is not plausible that this is the phenomenal content of John’s experience. But this is not enough to claim that the content is essentially indexical, for the skeptic also holds on PORTABLE. In order to decide if such a content can be read as the skeptic of essential indexicality maintains (GLOBAL and PORTABLE), or if it is *perspectival* content (PORTABLE and NOT-GLOBAL), we need to analyze whether such a phenomenal content is GLOBAL.

The content is GLOBAL only if it is assigned the same truth value for any point of evaluation within a world; that is, once we fix the way the world is we can assign the content a truth value. In this example, where the indexical expression is a first-person pronoun, the relevant point of evaluation is the subject of experience. Therefore, to evaluate whether the content is global we need to consider whether two subjects inhabiting the same world can

11 Please note that the argument in this section requires an even weaker premise, in which the entailment is restricted to cognitive systems like ours or actual humans. However, given that the skeptic is happy to accept even Brogaard’s stronger formulation of weak representationalism it is not necessary to press the idea any further—for discussion see Sebastián (MS)

have representations with the same content but different truth value. The answer to this question is undoubtedly positive for anyone accepting that veridical experiences and hallucinations or illusions can have the same phenomenology, something widely accepted.¹² Now, let Berit's experience be veridical but let John's be an illusion. Berit's and John's experiences have the same phenomenology and thereby the same phenomenal content. However, there is such-and-such sound coming from the left in the case of Berit but not so in the case of John. Since Berit and John inhabit the same world, and their experiences have the same content, fixing the world is not sufficient to assign their experiences a truth value. Phenomenal content cannot then be GLOBAL, for it assigns different truth values to different individuals within the world.

Summarizing, if different subjects can have experiences with the same phenomenology and these experiences share the phenomenal content, then phenomenal content must be PORTABLE. Moreover, this content cannot be GLOBAL to the extent that we allow for illusory experiences with the same phenomenology as veridical ones: its truth value changes within a world. Therefore, the phenomenal content of this kind of experience is perspectival: PORTABLE and NOT-GLOBAL. We can now come back to C&D's objections.

4. Back to the Skeptic Reply

As we have seen, C&D's strategy to resist the need of essentially indexical content in perceptual experience comes in two steps. First, they resist the arguments for indexicalized truth, that is, the rejection of GLOBAL. Second, they argue that even if one rejects GLOBAL this is insufficient for having essentially indexical content.

Undoubtedly, C&D can resist the rejection of global in the argument above in multiple ways. For example, they can deny that there are conscious states—states having phenomenology—, that conscious experiences have content, the entailment between phenomenology and phenomenal content or deny that veridical experiences and illusions or hallucinations have content. But this would miss the point of the discussion, and indeed C&D do not play that card in their reply. The aim of the argument above is not to show that there are essentially indexical contents but rather that there are philosophically interesting projects that require essentially indexical content and depart from classical theories of content. Sure, whether there is such a thing as essential indexical content then depends on whether any of these philosophical projects—indeed the vast majority in the study of consciousness and the mainstream view in perception—is on the right track. Years of discussion have taught us that this is not something to be decided

12 Sure, this idea is widely accepted but not universally so—for rejection see e.g. Martin (1998, 2002); c.f. Burge (2005), Dorsch (2010). Fortunately for current purposes it is not

in the length of a paper or a book. But this is fortunately something irrelevant for current purposes.

It is nonetheless worth considering something along the lines of the reply that C&D offer to Schroeder and Caplan (2007), and thinking of the possibility of phenomenal content not being a carrier of veridicality. This would result in a content that is not accommodated by traditional theories of content, which take contents to be the kind of entities that have truth value, but it is not clear that this tells us in favour of essential indexicality. Experiences with the same phenomenology would accordingly share a lambda-extracted content; in our example the property of *being an x such that there is such-and-such sound coming from the left of x* (denoted by: λx . *there is such-and-such sound coming from the left of x*), which they claim has no truth conditions. The suggestion seems to be that only particular experiences can be assigned a truth value but not so to the type they belong to in virtue of having a common phenomenology (the phenomenological type). Berit's experience predicates the property of *being someone such that there is such-and-such sound coming from the left* of Berit and it is veridical if Berit instantiates such a property, whereas John's experience predicates the property picked up by the lambda predicate of John and it is veridical if John instantiates such a property. The problem with this reply is that the lambda extracted content cannot play the role that the weak representationalist requires.

Experiences with the same phenomenology convey some common information regardless of the particular experience; they all restrict the uncertainties regarding possible alternatives, and they can be assigned a truth value depending on whether those alternatives correspond to reality. For example, if anyone has a token experience with the same phenomenology as Berit's, but there is no sound, then their experience will be inadequate. This gives us reasons to think that the content associated with the (phenomenological) kind—what I have called 'phenomenal content'—is indeed a carrier of veridicality. And of course, talk in terms of properties does not prevent this possibility, because property-content impose conditions on the entity of which the property can be adequately predicated and it is thereby informative in the required sense—see e.g. Tye (2019, 2021). Moreover, we should note that any token experience can only be adequate if the kind of property that C&D propose is instantiated by the subject of the experience. That is, Berit's experience cannot be veridical in virtue of someone else, for example John, *being an x such that there is a sound coming from x left*, because what is relevant for the veridicality of Berit's experience is whether Berit instantiates the property, not John. In general, the fact that someone S instantiates the property cannot make veridical a token experience, E^* , with same phenomenology as Berit's—regardless of the particular token experience we consider and whoever happens to be its subject—, unless S is also the subject of E^* .

A possible suggestion to solve the problem is to introduce, into the phenomenal content, explicit reference to the fact that the property must be instantiated by the subject of experience. For example, one might propose the phenomenal content of Berit's experience, E, to be the property of *being an x such that that x is the subject of E and there is such-and-such sound coming from x 's left* ($\lambda x. x$ is the subject of E and there is such-and-such sound coming from x 's left). If this were the phenomenal content of Berit's experience, then E, as we just demanded, would not be veridical in virtue of John being an x such that there is *such-and-such sound coming from x 's left*, because John is not the subject of E. But this move does not help to solve the problems. Recall that, if John has an experience (E') with the same phenomenology as Berit's experience (E), then E and E' must share the phenomenal content. And the property proposed is not a plausible content of E', because John is not the subject of E. John's experience is veridical if there is such-and-such sound coming from his left, not if there is such-and-such sound coming from the left of Berit (who is the subject of E).

Another possible suggestion is to replace the particular (E) by the phenomenological kind to which both, E and E', belong in virtue of having the same phenomenology; for example letting the content be the property of *being an x such that x is the subject of and experience with E's phenomenology and there being such-and-such sound coming from x 's left* ($\lambda x. x$ is the subject of an experience with E's phenomenology and there is such-and-such sound coming from x 's left). But this is of no help either. The reason is that both E and E' have the same phenomenology, but John's experience (E') cannot be adequate if he is hallucinating but Berit also has the experience and instantiates the corresponding property. In these circumstances, there is indeed such-and-such sound coming from the left of someone (Berit) who is the subject of an experience with E's phenomenology (E itself), but John's experience is inadequate (not veridical).

If we want to capture the informativeness of experiences with the same phenomenology, we will end up with a view that collapses into the portabilist one (Sebastián, MS). According to C&D, experiences with the same phenomenology have a common lambda-extracted content, which is a predicate that denotes a property. Experiences cannot be veridical unless it is the subject of the experience the one that instantiates the corresponding property—in the case of experiences with the same phenomenology as E, they cannot be veridical unless the subject of the experience instantiates the property of *being an x such that there is such-and-such sound coming from x 's left*. This, in turn, imposes a unique relation between the subject of the experience and such a property. The properties suggested by C&D to be phenomenal "content" are precisely the kind of properties that some portabilists such as Lewis postulate to be the content of *de se* representations. Lewis agrees that the content of a *de se* representation is a property, which

is precisely the one that results from the lambda-extraction. According to Lewis, *de se* representation involves a distinctive relation between a subject and properties. He calls such a unique relation in the case of belief ‘self-attribution.’ And as we have seen, talk in terms of properties being the content of attitudes can be directly translated into the canonical framework, holding on the classical relation that is determined by the attitude to a content that is evaluable for truth. All that is required, as Lewis points out, is to let the contents be centered propositions, where centered propositions are not classes of possible worlds but classes of centered worlds. Before moving forward, it is important to note that the appeal to centered worlds is a mere tool to translate a characterization of the content in terms of properties to the canonical framework. Hence, I am not relying, *pace* C&D, on speaking in terms of centered worlds to argue for essential indexicality. And this brings me to their second line of defense.

C&D dispute that talk in terms of centered worlds mandates an Essential Indexical content in chapter 5. More precisely, they argue that even if indexicalized truth is created, if GLOBAL is rejected, this does not amount to the creation of essential indexicality. They return to this point in the chapter devoted to perception (ch. 8), where they present the example of Maeve, whose perceptual system is fed by Brigid’s environment. Without offering any further justification, C&D (p. 145) assume that, in these circumstances, Maeve’s experience would be veridical if there is such-and-such sound coming from Brigid’s left rather than from her left, and hence, that phenomenal content does not involve self-attribution. However, this assumption seems to rely on a misunderstanding of the relation played by the actual cause of a representation in determining its content (Millikan 2000). Imagine that John wears light-filtering glasses, so that, when he is looking at a red object, he has the (phenomenological) kind of experience that we have when looking at green objects. In these circumstances, a red object is the actual cause of John’s experience, but red does not enter the phenomenal content of John’s experience, and his experience is indeed misrepresenting the color of the object he is looking at. Similarly, in the case of Maeve. Despite the fact that Brigid’s environment is the actual cause of Maeve’s perceptual experience, it seems more plausible to claim that her experience misrepresents the environment as her own—regardless of the belief state she is disposed to form on its basis. Assessing whether Maeve is indeed misrepresenting or not depends on one’s preferred theory of mental representation. But my point here is that there are *philosophically interesting projects*—for example those that combine weak representationalism with etiologial theories of mental content (Millikan 1984, 1989)—according to which Maeve is misrepresenting and hence immune to C&D’s remark.

We can set this issue aside, because C&D’s objection does not fundamentally rely on whether self-attribution is required, but rather, as they

stress, on whether phenomenal content can be interpreted non-indexically: “Maeve can have perceptual contents that have no absolute truth conditions, but whose role in Maeve’s perceptual system is to be true relative to Brigid” (p.149). C&D’s point in the end is that indexicalized truth, the rejection of GLOBAL, is not sufficient for essential indexicality. The content they propose might be NOT-GLOBAL but it is not *perspectival* because it is referentially anchored (to Brigid in this case). But phenomenal content cannot be such a referentially anchored content as we have seen. In the (phenomenological) kind of experience we are considering, Brigid definitely is not part of the phenomenal content of the experience of someone who is not Brigid, or whose perceptual system is not fed by Brigid’s environment—if we grant the content attribution that C&D make to Maeve’s experience—, and hence, not part of the phenomenal content of Brigid’s either. If John can have experiences with the same phenomenology as Maeve’s and Brigid’s, then his experiences have the same phenomenal content as theirs. Therefore, such a common phenomenal content cannot be referentially anchored to Brigid. Phenomenal content is expressed by deploying an indexical expression, which must be understood in the perspectival way: as PORTABLE and NOT-GLOBAL.

The argument discussed in the previous section offers, then, independent reasons, as those rejected by C&D, to accept portable. C&D deny, in chapter 5, that different individuals sincerely expressing the content of their beliefs by means of the same indexical expression have the same content. They deny that, when Hume and Heimson sincerely express the content of their beliefs by uttering ‘I am Hume’, they have beliefs with the same content, by rejecting Lewis’ internalism. The argument in the previous section offers independent reasons for phenomenal content being shared by different individuals when they use the same indexical expression to sincerely express it—for phenomenal content being portable.

To be completely clear, C&D claim, and I am happy to concede, that essential indexicality requires distinguishing the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were true relative to Maeve and the claim that Maeve had indexed-truth perceptual states whose veridicality conditions were truth relative to herself. Perspectival content makes this distinction. The phenomenal content of Maeve’s experience cannot be an indexed-truth perceptual state whose veridicality conditions were true relative to Maeve if John can have an experience with the same phenomenal character as Maeve. In such a case, John’s experience would have the same content as Maeve. However, his experience would not be veridical if it happens to be such-and-such sound to the left of Maeve, but not so to the left of John. Perspectival content has veridicality conditions relative to the subject rather than relative to any particular individual.

5. Conclusion

Cappelen and Dever (2013) claim that indexicality is not essential. They argue that classical considerations derived from the work of Perry and Lewis do not support the idea that we need to change our theories of content, and that essential indexical representations do not play any philosophically relevant role. In chapter 8, they note that if there are essential indexical thoughts, we should expect an essential indexical element in the content of perception, since those representations function as input of other thoughts. They acknowledge that perceptual content is typically expressed deploying an indexical expression but they argue that this content can perfectly be captured in non-indexical terms.

In this paper I have argued that traditional theories of mental content fail to characterize phenomenal content, because if experiences of different subjects can have the same phenomenology and thus share the phenomenal content, then such a content cannot be global and it must be portable. Phenomenal content is then perspectival content, which is not referentially-anchored with regard to the indexical expression deployed to express it. In other words, the entity that would be taken to be the reference of the indexical expression cannot be part of the phenomenal content. The argument in this paper attends Cappelen and Dever's complaints and connects the fact that perceiving subjects cannot be placed in the phenomenal content with what it takes to be represented in a distinctly indexical way, as it happens in *perspectival* (NOT-PORTABLE and GLOBAL) *de se* content.¹³

References

- Boer, S. E., & Lycan, W. G. (1980). Who, Me? *Philosophical Review* 89(3), pp. 427–466.
- Brogaard, B. (2009). Perceptual Content and Monadic Truth: On Cappelen and Howthorne Relativism and Monadic Truth. *Philosophical Books* 50(4), pp.213–226.
- Burge, T. (2005). Disjunctivism and Perceptual Psychology. *Philosophical Topics* 33, pp. 1–78.
- Byrne, A. (2009). Experience and Content. *Philosophical Quarterly* 59, pp. 429–451.

13 I am grateful to Ekain Garmendia and Dan López de Sa for detailed discussion on the *de se* phenomenon that has contributed greatly to the present paper.

Financial support for this project has been provided by DGAPA projects IN400520 and the Programa de Apoyos para la Superación del Personal Académico de la UNAM (PASPA).

- Cappelen, H., & Dever, J. (2013). *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*. Oxford: Oxford University Press.
- Cappelen, H. & Hawthorne, J. (2009). *Relativism and Monadic Truth*. Oxford: Oxford University Press.
- Carruthers, P. (2000). *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.
- Chalmers, D. (2006). Perception and the Fall from Eden. In Gendler, T., & Hawthorne, J. (eds), *Perceptual Experience*. Oxford: Oxford University Press, pp. 49–125.
- Chalmers, D. J. (2004). The representational character of experience. In Leiter, B. (ed.), *The Future for Philosophy*. Oxford University Press, pp. 153–181.
- Devitt, M. (2013). The myth of the problematic de se. In Feit, N., & Capone, A. (eds), *Attitudes de Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, pp. 133–162.
- Dorsch, F. (2010). Transparency and Imagining Seeing. *Philosophical Explorations* 13(3), pp. 173–200.
- Douven, I. (2013). The epistemology of de se beliefs. In Feit, N., & Capone, A. (eds), *Attitudes de Se: Linguistics, Epistemology, Metaphysics*. CSLI Publications, pp. 273–289.
- Dretske, F. (1995). *Naturalizing the Mind*. MIT Press.
- Fodor, J. (1987). *Psychosemantics*. Cambridge, MA: MIT/Bradford.
- Frege, G. (1956). The Thought: A Logical Inquiry. *Mind* 65(259), 287–311.
- García-Carpintero, M. (2017). The Philosophical Significance of the De Se. *Inquiry: An Interdisciplinary Journal of Philosophy* 60 (3), pp. 253–276.
- Guillot, M., & García-Carpintero, M. (forthcoming). *The Sense of Mine*. Oxford: Oxford University Press.
- Horgan, T., & Tierson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In Chalmers, D. (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press.
- Kölbel, M. (2013). The Conversational Role of Centered Contents. *Inquiry* 56(2–3), pp. 97–121.
- Kriegel, U. (2002). Phenomenal Content. *Erkenntnis* 57, pp. 175–198.
- Kriegel, U. (2009). *Subjective Consciousness: A Self-Representational Theory*. Oxford University Press, USA.
- Lewis, D. (1979). Attitudes De Dicto and De Se. *Philosophical Review* 88(4), pp. 513–543.

- Magidor, O. (2015). The Myth of the De Se. *Philosophical Perspectives* 29(1), pp. 249–283.
- Martin, M. G. F. (1998). Setting Things Before the Mind. In O’Hear, A. (ed.), *Current Issues in Philosophy of Mind*. Oxford University Press.
- Martin, M. G. F. (2002). The Transparency of Experience. *Mind and Language* 4(4), pp. 376–425.
- Millikan, R. (2000). *On Clear and Confused Ideas: An Essay about Substance Concepts*. Cambridge: Cambridge University Press.
- Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. MIT Press.
- Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy* 86, pp. 281–97.
- Morgan, D., & Salje, L. (2020). First-Person Thought. *Analysis* 80(1), pp. 148–163.
- Nagel, T. (1974/2002). What is It Like to be a Bat? In Chalmers, D. (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press.
- Nanay, B. (2014). Empirical Problems with Anti-Representationalism. In Brogaard, B. (ed.), *Does Perception Have Content?* Oxford University Press, pp. 39–50.
- Ninan, D. (2012). Counterfactual Attitudes and Multi-Centered Worlds. *Semantics and Pragmatics* 5 (5):1–57.
- Ninan, D. (2016). What is the Problem of De Se Attitudes? In García-Carpintero, M., & Torre, S. (eds), *About Oneself*. Oxford: Oxford University Press.
- Pautz, A. (2010). Why Explain Visual Experience in Terms of Content? In Nanay, B. (ed.), *Perceiving the World: New Essays on Perception*. New York: Oxford University Press.
- Perry, J. (1979). The Problem of the Essential Indexical. *Noûs* 13, pp. 3–21.
- Perry, J. (2001a). *Knowledge, Possibility and Consciousness*. Cambridge, MIT Press.
- Perry, J. (2001b). *Reference and Reflexivity*. CSLI Publications, Stanford.
- Rosenthal, D. (2005). *Consciousness and Mind*. Oxford University Press.
- Schellenberg, S. (2018). *The Unity of Perception: Content, Consciousness, Evidence*. Oxford: Oxford University Press.
- Schlick, M. (1959). Positivism and Realism. In Ayer, A. J. (ed.), *Logical Positivism*. New York: Free Press.
- Schroder, T., & Caplan, B. (2007). On the Content of Experience. *Philosophy and Phenomenological Research* 75(3), pp. 590–601.

- Sebastián, M. Á. (2012). Experiential Awareness: Do you Prefer „It“ to „Me“? *Philosophical Topics* 40(2), pp. 155–177.
- Sebastián, M. Á. (2022). First-Person Perspective in Experience: Perspectival De Se Representation as an Explanation of the Delimitation Problem. *Erkenntnis*.
- Sebastián, M. Á. MS: *Consciousness and Perspectival De Se Content*.
- Shaw, J. R. (2020). De Se Exceptionalism and Frege Puzzles. *Ergo: An Open Access Journal of Philosophy* 6, pp. 1057–1086.
- Shoemaker, S. (1982). The Inverted Spectrum. *Journal of Philosophy* 79, pp. 357–381.
- Shoemaker, S. (1996). *Intersubjective/Intersubjective, The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press, pp. 141–154.
- Siegel, S. (2010). *The Contents of Visual Experience*. New York: Oxford University Press.
- Siegel, S. (2011). The Content of Perception. In Zalta, E. (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition).
- Siewert, C. (1998). *The Significance of Consciousness*. Princeton University Press.
- Spencer, C. (2007). Is There a Problem of the Essential Indexical? In O'Rourke, M., & Washington, C. (eds), *Essays on the Philosophy of John Perry*. MIT Press, pp. 179–198.
- Stalnaker, R. C. (2000). Comparing Qualia Across Persons. *Philosophical Topics* 26, pp. 385–405.
- Torre, S. (2018). In Defense of De Se Content. *Philosophy and Phenomenological Research* 97(1), pp. 172–189.
- Travis, C. (2004) The Silence of the Senses. *Mind* 113(449), pp. 57–94.
- Tye, M. (1997). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. The MIT Press.
- Tye, M. (2002). *Consciousness, Color, and Content*. The MIT Press.
- Tye, M. (2019). *The philosophy of Perception. Proceedings of the 40th International Ludwig Wittgenstein Symposium. Publications of the Austrian Wittgenstein society*, chapter: How to Think About the Representational Content of Visual Experience, pp. 77–94.
- Tye, M. (2021). *Vagueness and the Evolution of Consciousness: Through the Looking Glass*. Oxford University Press.



THE ESSENTIAL INDEXICAL AND SELF-CONSCIOUSNESS: 'I', 'NOW', AND 'HERE' AS ASPECTS OF SELF-CONSCIOUSNESS

Abstract: *This paper aims to analyse egocentric indexicals 'I', 'now', and 'here' as different aspects of the same self-conscious or self-referential act emphasising the underlying phenomenological structure of the essential indexical 'I'. What makes an indexical essential is not its indexicality but the egocentric mental state indicated by its use. Therefore, interpreting them only in the confines of language severely limits the scope of the investigation. First, I will define the pure use of 'here', 'now', and 'I', which will consequently lead to the relation between the indexical 'I' and the pure I, and the transcendental designation of the subject. In the second part of the paper, the focus will shift towards the phenomenological notions of primal I with its contextualisation and to the dimensions of 'now' and 'here', analysed through the character of *nunc stans*. Through some ideas of Husserl and Heidegger, 'here' and 'now' will emerge as amalgamated with the I. It will be shown that they acquire their formal meaning, along with the empirical content, through the I, which is the actual I of immediate presence that is purely self-referring, and that other indexicals and pronouns are thus derivative. It will be concluded that the essentiality of the indexical 'I' has its origin in the very structure of pure self-consciousness. In that regard, the theory of the essential indexical 'I' could be interpreted as an extension of the doctrine of the pure or transcendental I.*

Keywords: *the I; 'now'; 'here'; self-consciousness; indexicality; phenomenology.*

“This, 'here' and 'now' cannot be learned or understood except by one who has general, acquired awareness about him- or herself”
(Frank 1999, 207–208)

“A self is what is referred to in the first person: what the first-person pronoun denotes in its proper and correct use. Hence, the structure of the mechanisms of self-reference reveal at least part of the structure of the self.”
(Castañeda 1981, 62)

What makes an indexical *essential*? According to some, it is their ineliminability and irreplaceability – the speaker's belief cannot be fully expressed in language that does not contain indexicals. Those are locating

beliefs, i.e., beliefs about “where one is, when it is, and who one is” (Perry 2001, 145). They are essential in the sense that replacing them with a nonindexical term destroys the explanation or requires further assumptions. If we replace ‘I’ with another designation of me, we no longer have attribution of the same belief. I shall call this description a third-person description, one which can be given by any individual about any object and where the meaning is preserved. Perry’s statements “Someone is making a mess” or “John is making a mess” belong to this category, while “I am making a mess” would belong to a first-person description. ‘I’, when I utter it, cannot be replaced by a third-person description in such a way that I recognise *myself* in that description as the subject (i.e., as John or as ‘someone’).¹ In other words, the “translation” from third-person to first-person is not without a loss in meaning or ambivalence. It cannot be replaced even with a complete description or a thing (Shoemaker 2001). They are also necessary for action, as shown in the example of the careless shopper (Perry 2001) or being attacked by a bear (Perry 1977), and also in Lewis’ example of two gods (Lewis 1979).

This view, however, has been challenged. It is not their indexicality that makes them essential, but the structure of egocentric mental states and first-person redundancy (Prosser 2015). Alternatively, in Millikan’s words, “[it] is not their semantics that distinguishes [indexicals], but their function, their psychological role” (Millikan 2001, 163). In this paper, I shall be leaning towards this interpretation and offer my take on the idea that the theory of the essential indexical ‘I’ could be interpreted as an extension of the doctrine of the pure or transcendental I.² Interpreting ‘I’ only in the confines of language severely limits the scope of the investigation. I argue that ‘I’ and the I are both essentially and structurally tied together with the first-person perspective and that all other indexicals formally refer to or have their content determined in relation to the I, whilst it transcendently designates the subject of thought (Brook 2001). I shall also be focusing on what I believe are three main egocentric indexicals, ‘I’, ‘now’, and ‘here’. Firstly, what I term a “pure” use of ‘now’, ‘here’ and ‘I’ will lead to the doctrine of the pure I, after which the relation between ‘I’ and the I will be sketched. In the second part, I will analyse themes from phenomenology, such as notions of *nunc stans* and primal I.

1 One could argue that it is *vice versa*. I can recognise myself in a third-person description of myself based on some implicit cognitive elements and immediate acquaintance with myself. If someone says, “John made this mess”, and I know that my name is John and am already aware in the indexical form that I made that mess, then I can be fairly certain that he is talking about me. However, the “jump” from third-person to first-person, i.e., nonindexical to indexical, is more problematic, as shown in the case of Rudolf Lingens (Perry 1977).

2 I shall be using the term ‘I’ when discussing the indexical, focusing on the word, expression, or a first-person pronoun, and ‘the I’ to refer to the pure or transcendental I or ego.

1. The pure use of ‘now’ and ‘here’

Under the term ‘pure’, I understand in the widest sense that which is free of experience or can have experiential or sensible content. To that end, the “purity” will be shown to be a mode of self-referentiality.³ It could also be followed along the lines of Kaplan’s distinction between linguistic meaning or character and the content of an indexical; or Millikan’s “indexical adapting relation” (Millikan 2001, 165). Roughly, the main difference between pure indexicals and other demonstratives is that the former have their content determined by their meaning. ‘I’ has a single meaning, it “refers to the speaker or writer” (Kaplan 1989, 505) or “the producer of the token” (Millikan 2001, 165), but it can have different contents depending on the context. ‘I’ refers to the subject uttering it, and it cannot *not* refer to it, or refer to someone or something else. On the other hand, demonstratives often require additional information or cognitive element in order to know what the subject is referring to; they merely constrain semantic reference. If I point my finger at a chair and say “This”, it (hopefully) refers to the same thing for you as it does for me.⁴ By contrast, ‘I’, ‘now’ and ‘here’ are self-oriented and from the first-person perspective. Thus, their content is dependent on the same self-evidence that Descartes employed in his *Meditations*, which is to say, it is undoubtedly true that ‘I am’, as it is that ‘I am now’ or ‘I am here’. We could try and take the opposite to be true by saying “I am not here” or “I am not now” and see that it is formally impossible. Interestingly, it is precisely the indexicals for context’s speaker, time, and location, i.e. ‘I’, ‘now’, and ‘here’, that have *automatically fixed references*.

There are cases of ‘now’ and ‘here’ without content, i.e., I am able to say ‘here’ without explicitly thinking of any location. If I am in a search group and find something, I can yell to others, “Over here!” or simply “Here!” without looking where I am or orienting myself. I do not need to know where I am in order to call someone to me. My utterance does not contain “The place between the hollowed oak tree and the ditch”, etc. It simply means ‘Come to me’ or along the lines of ‘Follow my voice’, ‘I want you to come to the immediate presence of me’ to stand next to me, or ‘I have found *x* at the location from which I am yelling’. At any moment, I can look around myself and offer a description “I am in between the hollowed oak tree and the ditch”. However, that only comes after the reflection and “empirically” determining

3 I understand self-referentiality in the broadest sense of something making a reference to itself as it is.

4 This is also problematic, for I can say, “This is a strange situation we find ourselves in”, referring to the very context, or “That is a good point”, referring to the previous argument, without any pointing or visualisation, compare (Millikan 2001, 169). Or, perhaps, our “field” of thought or consciousness carries in it some remnants of structural spatiality, like the one found in locating an event in time (i.e., representing time and events as an *x*-axis) or having one thought “next” to another. But this is a topic for another paper.

the location in the relevant context or current situation. Formally, I am determining the location as *my immediate surrounding* or, more precisely, that the location where I found *x* is identical to the location where I am (now). Needless to say, that is a tautology and does not say anything in detail. Still, it does not have to, because it will have all the meaning to someone who heard me. They would know where to go (to the direction in which they heard my voice) and would not need any more descriptive or precisely detailed account of the location. In a report I give the day after, I would have to say, “I have found *x* next to the hollowed oak in the ditch.” I can’t say that I found *x* ‘there’ or ‘here’. The only way I could use ‘here’ afterwards is if I took the detectives to the same place and said, “I found *x* here”. ‘Here’ does not tell us *where* it is and accordingly is of no use uttered in a dark room to a person with one deaf ear who cannot localise sounds (Millikan 2001, 168). Even so (as I will soon emphasise), it does formally tell us about the location based on self-referentiality. In other words, the only reason why anyone would follow the sound of “Here!” is because they are aware that ‘here’ in “Here!” refers to the place of origin of the sound. What differentiates the pure use of ‘here’ from the aforementioned accounts is precisely its *use*; it is not just a linguistic convention or rule for it.

As Prosser formulates, “[if] *S* refers to a place *l* using the word ‘here’, then normally *S* believes that *l* is where *S* is located. *l* is, as *S* might put it, *hereabouts*.” (Prosser 2015, 215). Both Smith and Jones could believe that there is danger at *l*, where Smith is, but “Smith also believes that *l* is hereabouts, which is true if, and only if, *l* is where Smith is located”. While I agree with Prosser’s insight, I do not think that the ordering is correct. Smith does not believe that *l* is ‘here’, or hereabouts for him, but ‘here’ is *l* in that particular occasion. In other words, ‘here’ is not a property of *l*, or that the empirical content of ‘here’ is firstly thought, instead I would believe that there *is* danger (here), because it is in my immediate presence (or vicinity), and for practical reasons of conveying information to someone, I would describe ‘here’ as *l*. I am not where ‘here’ is; instead, ‘here’ is where I am. More so, ‘here’ is a self-reference to *my* location; *l* just happens to be that location. It follows ‘*T*’ as the circumference of a circle follows its centre. ‘Here’ could also be interpreted as the “presentness” of the *I*, in a sense that it becomes (contentwise) that location which ‘*T*’ “illuminates” by being present.⁵

The same holds for ‘now’. ‘Now’ does not just so happen to be the time *when* I am, i.e., I am not when ‘now’ is; instead, ‘now’ is when I am. If this

5 English language has that peculiarity that the term ‘present’ can have both *temporal* meaning, that something is now occurring; and *spatial* meaning, but not in a straightforwardly locating sense, i.e., denoting a place, but in a case of more abstract presence or absence, that something is at hand, attending or simply here. For example, a doctor has to be present during the operation, but it is not specified where exactly he has to be – if it is in the same room or if he, in some futuristic setting, could act through a robot or oversee through a video call from the other side of the country. He just has to attend.

were not the case, we would have a colossal task of explaining why everyone, every 'I', is at the same time, i.e., in the same 'now'. We could disagree on what time exactly 'now' is, I could strongly believe that it must have passed noon by now, and you might argue that it is probably still around 11 am. Nevertheless, we will never disagree that 'now' is precisely now, as we are speaking about it. Our "inner clocks" do not have to be synchronised, but my 'now' cannot be a couple of minutes (or some "time") behind or ahead of yours.⁶ 'Now' has nothing to do with time but with self-referentiality.

In both cases of 'here' and 'now', in the same way as "indexicals do not *tell* what they point at [...] what is in their contexts" (Millikan 2001, 169), I argue that *nothing in the content tells about the indexical*. Nothing in *l* contains that it is 'here', or that some date and time are 'now', or that John or Rudolf Lingens is me. 'Here' is not a property of a location. There is nothing in my office (or at *l*) that makes it 'here' – my office was 'here' only because I was there. Spatial and temporal self-localization is an ability independent of any concrete location or time; it follows the I or self-reference and one's "'I'-Idea and one's 'here'-Ideas are really two sides of a single capacity, each wholly dependent upon the other." (Evans 2001, 140).⁷

Still, they are more than simple linguistic terms or what is expressed by egocentric mental states as *formal vehicles of the first-person perspective, self-conscious subject*. The pure use demonstrates that indexicals 'now' and 'here' are "placeholders" in the self-conscious mental act (Castañeda 2001). In essence, I propose that 'here' does not refer to some (empirical) *l*, but to my hereabouts, my Here, the presentness of the subject, whatever that location might be, in the Now. It is contingent that 'here' at t_1 (or now_1) is my office and at t_2 (or now_2) my home. Their content is not just fixed by their meaning, but also, formally and empirically, they relate to the I. Their meaning is independent of any context (or content), and their formal or pure use is also without one. In this sense, they are self-referring and indicative of an objective, formal structure of mental states or consciousness, irrelevant of the content.

Let us now turn to the pure use of 'I'.

2. The pure use of 'I'

Suppose we differentiate between 'I' when I use it and when others use it. In that case, for others, I can know for certain that formally it refers to the subject of utterance, but materially (contentwise), I do not need to know *who* it is. Imagine the following scenario. A night security guard is on the

6 This strain inevitably leads to the metaphysics of time.

7 Self-localization is the ability to locate oneself in the environment, to identify the immediate surroundings. Immediate surroundings are always (formally) identified: it is the space or place around *me*. If I wake up in a strange place, I will try to identify *where* 'here' is.

watch when he hears some noise. He yells, “Who goes there?!” and hears me reply, “I!”. If the guard recognises my voice and knows me, he can know my identity. However, if he does not know me, the word ‘I’ bears no other meaning for him other than that it is indeed a person, i.e., a subject that can meaningfully use the indexical ‘I’, and not some wild animal that is making the noise. However, from my first-person perspective, I have told the guard everything. I have identified the person who is trespassing by admitting that it was indeed me and not someone else. The difference between this kind of admitting and, e.g., if the schoolmaster asks, “Who broke the window?” is that in the second example, the schoolmaster can see me uttering the word ‘I’, whilst the guard only hears the voice coming from the darkness. ‘I’, in the case of the schoolmaster, means that it was not “him, or him, or her, or them (a gang of mischievous students)”, but that the perpetrator is the one speaking. Just the same as with ‘here’ and ‘now’, “a token of ‘I’ does not tell me who the originator of that token is” (Millikan 2001, 168), except, I add, in the first-person perspective.

We can differentiate between the indexical’s formal (pure) aspect and its (empirical) content. With ‘this’, I could refer to an unlimited number of things. Similarly holds for ‘now’ and ‘here’, albeit limited. If I say, “This is red”, you can, if you are not looking, be mistaken as to what necktie I am exactly referring to (Shoemaker 2001, 84). On the other hand, for ‘here’ you cannot be mistaken, although you have to see where I am or be able to orient yourself. If I hear a voice saying, “It is cold here”, and I cannot precisely determine if I heard it coming from the outside, through an open window, or from down the hallway, through an open door, I could only be sure that it is cold at the place from which the person in question is uttering the sentence. But this kind of error is impossible with ‘now’. You do not have to look at me, at what I am pointing or what exactly am I doing to know that ‘now’ refers to the very moment at which it is being uttered.

We can derive a gradation of certainty. For ‘now’, I only need to hear the sentence being uttered because there is only one current now.⁸ But for ‘here’, I also am to perceive the location of the subject that utters it or to be able to orient the direction from which the sound is coming (albeit, formally, I already know where it is, i.e., it is the whereabouts of the I uttering it). For ‘this’, absolute certainty is impossible. Consequently, I can choose the reference of ‘this’ (this table, this problem, etc.), and I can select the reference for ‘here’, but I cannot choose the reference for ‘now’ or ‘I’. I argue that ‘I’ falls in the first category, with ‘now’; they are permanently fixed.⁹

8 What is more, ‘now’ at which I am hearing it is the same ‘now’ as that of the utterance. My ‘now’ and your ‘now’ are the same ‘now’ if the utterance is actual; they are the same actuality. If someone says, “It is raining now”, it means that I am also going to get wet if I step outside. But if I hear a recording of the utterance or read it, I know that that ‘now’ is not the actual Now, ‘now’ of the act of hearing, which can never be a nonactual Now.

9 This is in agreement with Descartes’ evil demon. He can deceive me of where ‘here’ is or if ‘here’ even exists (if I am only a brain in a vat or a subject of solipsism). Virtual

Is ‘I’ a placeholder for *me*, i.e., this person I am? ‘I’ can only have one object that it is placeholder for. Opposed to every other indexical, the content (or referent) of ‘I’ is always identical – *it is me*. I can never use a different subject for ‘I’, nor can I designate any other subject different from the one *who is designating*. One day I could be standing in front of an audience in the amphitheatre and say, “I am here, now”, and the next day at home and say, “I am here, now”. ‘Here’ said in the amphitheatre means, contentwise, the amphitheatre or, formally (i.e., in relation to the I), the location where I am uttering it; and the same goes for my home – and I can choose it. ‘Now’ can also mean, contentwise, one time or date or the next one or, in relation to the I, the moment at which I am uttering it – but I cannot choose it. However, in both cases, the referent of ‘I’ is the same, unchangeable – *and I cannot choose it*. On both occasions, it means the same thing, *me*. What is more, if ‘here’ is where I am now, and ‘now’ is when I am uttering it, then ‘here’ is tertiary, for it depends on the ‘I’ and ‘now’. ‘Now’ is secondary because it only depends on the actuality of consciousness, i.e., ‘I’ or the moment of self-reference. And ‘I’ is primary, for it does not depend on anything; it simply is or is not. Self-reference is either occurring, or it is not, and if it is, then *when* it is occurring is ‘now’, and the result is unchangeable, it is ‘I’. There is no self-reference occurring in the past or future, one that is not actual. That kind of self-reference has happened or will happen. The self-reference that *is* occurring is the one that is “underway”, the actual one.¹⁰ The same can be said for existence, for there cannot be my existence outside of me actually existing.¹¹ If there is an actual I, there also is a ‘now’, and *vice versa*.

I contend that ‘I’, ‘now’, and ‘here’ are formal vehicles in the perspectival structure of consciousness. Their equivalents in language simply have indexical or demonstrative nature. It is not because of their indexical nature that terms like these influence the development of our consciousness, but *vice versa* – it is because they are formal elements “operating within” our consciousness that they have indexical nature.¹² This only means that ‘now’ and ‘here’, or their nonlinguistic mental equivalents (the Now and the Here, if we ought to term them in accordance with the I), have nothing empirical or

reality can already do this. Noteworthy is that even in that case, a talk of a presence or absence of something (to me) would still be valid. Yet, he could not deceive me that he is deceiving or that he is not deceiving me now, at my moment of doubting it or thinking about it; i.e., he cannot deceive me that ‘now’ is not exactly this moment. And, finally, he cannot deceive me that I am not the one who is being deceived.

10 This does not mean that we cannot talk about the notion of self-reference or the act in general. Nevertheless, if there is actual self-reference, it cannot be divorced from time, more precisely, from the Now.

11 See note no. 26. I believe that this kind of self-referentiality is what is nested deep in the *cogito* argument.

12 See “an *egocentric mental state*, which is essential for action, is indicated by the use of an indexical term” (Prosser 2015, 212). More on this later.

contextual tied to them. Consciousness operates on the level which is pure, i.e., abstract from any context or relevant occurrence. And if that is the case, then self-relation must also be pure, hence, the pure I. It would also imply that 'I' is not (just) an indexical or that the source of its essentiality lies within pure self-consciousness.

3. The I as an 'I': transcendental designation

How are we to understand the pure I? There is a range of interpretations in the continental tradition, but it will suffice to define it in Kantian and Husserlian terms.

I take myself as the pure Ego insofar as I take myself purely as that which, in perception, is directed to the perceived, in knowing to the known [...] there lies a ray of directedness [that] takes its point of departure in the 'Ego', which evidently thereby remains undivided and numerically identical while it lives in these manifold acts. (Husserl 2000, 103–104)

Suppose I am perceiving an object, thinking about x , etc., then I am in all these acts that which perceives, thinks. Being the pure subject means being defined only by the act in which one is found. Yet, the subject of different acts is not itself different or separate; there is no multitude of subjects. It is one and the same in different *cogitationes*. It is nothing outside of the mental act, and in it, it is intentionally directed toward the object. It could also be interpreted as a primitive but necessary centre of relations, a centre of unity of consciousness, an I-pole.¹³

For Kant, in the original synthetic unity of apperception, "I am conscious of myself not as I appear to myself, nor *as* I am in myself, but only that I am." (Kant 1998, 259). The consciousness of oneself is not the same as the cognition of oneself. In another place,

Through this I [...] nothing further is represented than a transcendental subject of thoughts = x , which is recognised only through the thoughts that are its predicates. (Kant 1998, 414)

This subject of thoughts is "designated only transcendently through the I that is appended to thoughts, without noting the least property of it, or cognising or knowing anything at all about it." (Kant 1998, 419). In this I, nothing manifold is given; it is simply an awareness of oneself as a subject without the need to know anything (empirical) about oneself. Brook, with whom I am much in agreement with, develops the idea of transcendental designation as an "awareness of oneself as oneself" (Brook 2001). It is not enough to be aware of oneself – I can do that and not realise that it is me

¹³ For Husserl's notion of the pure I see (Husserl 1983).

(as in the case of Rudolf Lingens), instead – one has to be aware of oneself *as* oneself. And *vice versa*, I can refer to myself not just as John or Rudolf Lingens, but as *myself*. Awareness of oneself *as* oneself is, thusly, tautological. Precisely what this ‘as myself’ is or means is, I believe, the pure I. This is the single common subject of several different mental states. It would seem that ‘I’ functions in the way in which the guard and schoolmaster hear it – it *purely or transcendently designates the subject of its utterance*, i.e., we do not have to attribute any other way of referring to oneself “aside from his ability to use the pronoun ‘I’ or his ability to be conscious of himself” (Castañeda 2001, 60).

Let us consider a follow-up example. The guard hits me on the head, and I lose consciousness. I wake up later in a hospital with complete amnesia. I cannot remember who I am, what had happened, or how I got there. If I say to the doctor, “I cannot remember anything”, I will be correctly referring to myself or self-referring as the-one-who-cannot-remember-anything, without any empirical content or knowledge other than my immediate surroundings and state of affairs (that it is true of me that I cannot remember). I will also know that I am referring to *myself* by saying ‘I’, even though I do not know *who* I myself am. The reason being that ‘who’ I am does not have anything to do with my self-consciousness and self-referring. Me being self-conscious makes me being a ‘who’ (this or that person) possible. For what can an amnesiac think under the ‘I’? – Surely, the answer would be that he thinks himself (he is not thinking the doctor or the patient next to him by uttering ‘I’); except, what empirical knowledge is contained under the label ‘himself’? It cannot be his personal name, a representation, or some kind of a mental image of himself, his preferences, face, and all the things that constitute personal identity. With ‘I’, he can only refer to himself and ‘himself’ is the I as an object for itself. In other words, ‘himself’ is ‘the one who is uttering (this word) ‘I’’, or these words, meaning the subject of ‘this here utterance’ or, more poetically, ‘the subject of this voice.’ ‘I’ refers purely to the subject of the utterance of the ‘I’, or I *objectify myself*.¹⁴

That would not be all of which I would be a subject, as it would also be true of me that I am in a hospital, lying in bed, or talking to the doctor, etc. I would also have my ‘here’ and ‘now’ without any content (especially if I cannot recognise where I am the moment I wake up). I could say to the doctor, “I, who am now speaking to you, the doctor, cannot remember anything”. However, these kinds of contextualisations would, in a sense, be analytic *a priori* judgments; even so, they would still convey meaning to *me*. And without any context, empirical or analytical, I would still “know” that I am me, myself. This tautology is necessary for self-consciousness. Just as ‘here’ means my immediate surroundings, my hereabouts, wherever it may

14 This also cannot be Wittgensteinian use of ‘I’ as subject, for it is not simply an expression as in “I have pain”.

be, 'I' means me, myself, whoever that may be. And because it lacks any empirical or sensible content, it is defined tautologically, i.e., "I am the one who does not know who he is".

Beyond this 'information' that the thing presented is me, the representation tells me nothing about myself. This barrenness in one's awareness of oneself as subject was perhaps one of the things that led Hume to think that no subject is to be found in self-awareness at all. (Brook 2001, 29)

Therefore, even though 'I' has no (empirical) content, it is not without the referent. The referent of 'I' cannot be contextual, for it would be different in different situations. Indexicals depend upon the context,¹⁵ changing with it. Indexical thought would imply that it has its content and referent determined in a context and that we have a way of identifying them. However, the first-person context never changes, nor is it transformative. This is precisely why 'I' cannot be simply an indexical. If in 'perceiving *x*' I am the one who is perceiving, then its content would always be dependent on its context. However, in all mental states, the I as pure I remains self-identical and numerically one, as Husserl and Kant stated. Yet, I do not require any identification procedure to identify myself *as* that subject, the subject of thought. It is not identified simply by the context. In which case 'I' could still be an indexical, but one which content for me does not vary in contexts – *essential*. Yet, the I as pure I is the I-orientation of every context and every situation (Husserl 1970, 171), (Husserl 2000, 111).¹⁶ The referent point will always be the pure I which follows all my representations. It is the I-centre towards which all synthesis is directed, and which can follow with Kantian 'I think' (Husserl 1977, 158ff). We experience the world from a first-person perspective. Intersubjectively, 'I' depends for its content on the subject uttering it (as with the guard), but subjectively, its meaning is fixed, "static". It will never signify any individual other than myself in my conscious mental life. In other words, indexicals are dependent on the perspective except 'I', which *is* that perspective or orientation. If the perspective changes, 'I' has to transform into 'you' or 'he'.¹⁷

We can now return to previous examples. 'Here' does not convey the place where it is, and 'I' does not say who is the originator or the speaker of 'I'. This is peculiar because (in the example with the guard), it does not say anything *to you*, but it says everything *to me* (in both examples of the amnesiac and the guard). Still, my pure I or my use of the indexical 'I' and your pure I or your use of 'I' have to be the same; only what we later think under this 'I' is entirely different (after all, there is a reason why we can understand each other by

15 But they do not tell anything *about* it (Millikan 2001).

16 Compare: "dare I call it ego-orientation?" (Kaplan 1989).

17 More on contextualisation and transformation of the I will be said in the next section.

saying the same word but thinking or meaning completely different things). There is nothing “personal” in ‘I’. Paradoxically, ‘I’ would be an indexical that is *only mine*; something private, opening the *Me* for me, and at the same time public and intersubjective, and meaningful to any range of self-conscious beings.¹⁸ In other words, my use of ‘I’ has to be accompanied by underlying self-relation, which is not present in your use of ‘I’. Accordingly, we can think of ourselves empirically (as John, this concrete person, etc.) and purely (as the I, the one who thinks). Here nothing is contained as ‘the one who thinks’ is identified transcendently. At the same time, there is something indexical in the I, that it can signify different subjects in different contexts, and something essential, that my ‘I’ purely designates me in my self-conscious act.

Furthermore, ‘I’ is not simply a linguistic term because it must be used. A machine can replicate the sound ‘I’, which will not have any meaning. Similarly, in the first example, the guard only hears ‘I’ being uttered (which could have very well been a recording placed there), and in the second, I do not hear ‘I’ being spoken (like I would from the room next door); instead, I am uttering it, I am using it. I am consciously self-referring, in which case I know about whom I am speaking. Proper use of ‘I’ implies self-consciousness or an egocentric mental state.

Let us now turn our attention to Husserl’s notion of the primal I, where these ideas are further developed in the light of contextualisation.

4. The primal I: contextualisation of the referent of ‘I’

For Husserl, the I can occupy a special place in opposition to the world and objects. When it achieves that through *epoché*, ‘you’, ‘we’, and the entirety of intersubjectivity becomes a phenomenon for me, “the whole distinction and ordering of the personal pronouns, has become a phenomenon within my epoché” (Husserl 1970, 184–185). The I, because it does not have a ‘you’ opposed to it, is an I only by *equivocation*. It is the primal I, the I of my *epoché*, which can never lose its uniqueness and personal indeclinability. Indeclinability in the sense of absence of change or transformation into other perspectives, i.e., I cannot abandon my first-person perspective.¹⁹

Living in the world, “I am necessarily an ‘I’ that has its ‘thou,’ its ‘we,’ its ‘you’ – the ‘I’ of the personal pronouns.” (Husserl 1970, 335–336). In other words, the use of ‘I’ in language or ‘I’ as a pronoun is based upon the

18 Which could, to some extent, be associated with Mead’s idea of internalising the structure of the Self (as with language and symbols), particularly the I-Me relation. If each Self had a different structure, we would not be able to interact with each other; and yet, the Self is something private and introspective, see (Mead 1972).

19 Differently, the nominative case ‘I’ refers to the subject, but accusative or objective ‘me’ (or any other) is the subject taken as an object. It is not a *subject proper* but a reference to the subject. The subject can never be in any other case than nominative and still be the subject.

I. This I then is the first-personal character of consciousness, absolute formal individuation of the subject's consciousness (Zahavi 2014, 84); a structural perspectivity that is an indeclinable and ineliminable aspect of consciousness. Therefore, the I is still not an 'I' among others (with 'you', 'we', etc.), "all such distinctions as 'I' and 'you', 'inside' and 'outside', first 'constitute' themselves in the absolute ego." (Husserl 1970, 82). This would mean that indexicals have meaning *for me* and that the I can still exist for itself, i.e., be self-conscious, with or without intersubjectivity.

The consequence is that the I does not depend on intersubjectivity for its original structure. First-person consciousness cannot be "learned" or "internalised" in socialisation. The I has its formal, pure structure (e.g., self-referentiality, which could be a purely logical function) but constitutes its concrete individuation in contexts. The I makes itself declinable by *entering into contextualisation*. It integrates itself into a context of intersubjective relations between I's. The human being develops in society, but the structure of its consciousness does not come into being. Taguchi interprets self-contextualization in Husserl's words as a "change in signification of [the form] 'I' – just as I am saying 'I' right now – into 'other I's,' into 'all of us,' we who are many 'I's,' and among whom I am but *one* 'I.'" (Husserl 1970, 182). This change in signification is a modification of the primal I.

We can again return to the example of an amnesiac. It could very well be Rudolf Lingens, who cannot remember who he is and is reading a complete biography about Rudolf Lingens (Perry 1977, 492). In this case, we can interpret Rudolf Lingens before amnesia as the contextualisation of his pure I – he is this, and not some other person, with this name, history, etc. After the accident, however, what is left is empty consciousness with preserved I, i.e., self-reference. The pure I contextualises itself into an ego.²⁰ The most important thing to note here is the fact that after being contextualised or modified,

[t]he 'primal I' does not dissolve into its modification. It still remains as 'primal I;' and the contextualized ego constantly refers back to the 'primal I' whose modification it is. In other words, the 'primal I' places itself in a context that is composed of a relationship with many others and finds itself in a contextualized form, but at the same time, it does not lose its fundamental trait. (Taguchi 2018, 36).

Rudolf Lingens used 'I' to refer to himself *as* Rudolf Lingens. In that use of the egocentric indexical, in the persona of Rudolf Lingens, there was embedded his pure I. After the amnesia, everything empirical got "removed", yet the I remained. More precisely, the persona of Rudolf Lingens was built upon and

20 For Taguchi's example, see (Taguchi 2018, 37). This interpretation in some views could be wrong because what Husserl thinks under this contextualisation, and especially habituation, is "deeper" into the structure of consciousness than the level of memories and personality. Still, it suffices for the needs of this paper.

sedimented around the core I. The contents of consciousness, his beliefs and knowledge about himself, were lost, but that it is still *he himself* (who cannot remember, who is laying in a hospital bed, etc.), is retained. If that aspect of self-consciousness were also lost – *he would not have woken up from the coma*. There would be no consciousness in that situation, not because of some lack in the content, but because the very processes of consciousness or the way in which it functions would have been disrupted.

Precisely this I, pure self-consciousness, that remained is indeclinable and ineliminable, original self-givenness or first-person perspective of consciousness. In every context or situation, it is the same. If in ‘perceiving *x*’, I refer to myself *as* the one who is perceiving, or as Rudolf Lingens in identification, etc., in all those cases, the relation I-myself or I-Me has to be maintained. In the contextual self, or ego, personal I, it is the I that is constantly referring back to itself as indeclinable, pure I. The same kind of self-reference occurred when Rudolf Lingens uttered “I am Rudolf Lingens” and when now, as an amnesiac, utters “I cannot remember who I am”. The I’s self-reference to itself *as itself* is context-independent. The case of the amnesiac is just an extreme case of retaining this ability. I am both a non-contextual primal I and an individual ego among others – *at the same time*, and I experience myself as such.

In Zahavi’s view, when Husserl speaks about this, he does not have in mind the metaphysical status of the I; rather he is “pointing to its indexical nature” (Zahavi 2015, 6). As he mentions further, this does not mean that Husserl speaks of or reduces the I to indexicality, and, in my opinion, goes to show that the use of ‘I’ is deeply rooted in the pure I, as a part of the structure of consciousness.

What is more, according to Husserl, in remembrance, “to what is recollected, what is past [...] belongs also a past ‘I’ of that present, whereas the actual, original ‘I’ is that of immediate presence [...] Thus the immediate ‘I’ [...] constitutes a variational mode of itself as existing (in the mode of having passed).” (Husserl 1970, 185). If I remember a fire, I do not simply remember an event or an object, ‘*a fire*’ (without a mode in which it was given to me) or a (nonegological) ‘*seeing a fire*’, rather (egological) ‘*I saw the fire*’ – I remember that *I* have witnessed it (Husserl 2019, 297).²¹ If this is the case, there is a kind of identification in every mental act with the subject of that state.²² In our examples, it would mean that ‘There is such a subject as Rudolf Lingens’ followed with ‘*I am Rudolf Lingens*’ and ‘There is such a subject who had witnessed the fire’ followed with ‘*I am that subject*’. If there has to be such a connection between ‘Rudolf Lingens’ and me, then we can assume that it is

21 Compare Shoemaker’s example of “I see a canary” and “I saw a canary” (Shoemaker 2001, 85).

22 For an insightful paper on the topic of I-splitting, and simultaneous differentiation and identification of the I with itself, see (Cavallaro 2020).

also necessary in the act of remembrance. In every instance of self-knowledge, there should be implicit self-identification, the same one as in the case of always-returning essential indexical 'I', i.e., if I replace "I am making a mess" with "John is making a mess", then I would believe that 'I am John', which brings me back at the start (Perry 2001, 144). Or, the iterations *in infinitum* of third-person propositions, until first-person redundant mental representation is reached (Prosser 2015, 227). Indexical statements are impossible without self-consciousness, or propositions with indexical elements are impossible without self-conscious entities. The functioning of the essential indexical 'I' would be retained in the very structure of consciousness after amnesia. If some futuristic treatment for amnesia would be to simply reinsert or download memories back into the brain after an accident, then they would have to be in the mode of 'I' or that I of the immediate presence, of the Now and Here, would have to be able to recognise (and identify) myself as their subject.

I don't say 'someone ate oatmeal for breakfast this morning,' but rather 'I ate oatmeal for breakfast this morning.' By using the first-person pronoun, I am affirming that the identity of the one who ate the oatmeal and the one who is now recalling the episode is phenomenologically given. (Zahavi 2021, 272).

In the next section, we will return to the relation of the I and the Now of the actual I.

5. The absolute actuality of the Now and self-referentiality

The I is a subject, but it cannot objectify itself if not to transform itself into an object for itself, i.e., the Me. In Mead's words, "the self cannot appear in consciousness as an 'I,' [...] it is always an object, i.e., a 'me'" (Mead 1913, 374ff). I am a subject, yet I think about myself as an object, thus ceasing to be the actual subject who is thinking.²³ The I gets transformed into the Me as it enters the experience or the past. I have my past, but I am not *in* my past as an actual I – in that past, there is a subject, and the subject of that past is *me*, and Me is an objectified I. Nevertheless, the I is always "in the now"; at this moment, it always remains actual. It is the subject, 'I' of 'now' and 'here'. Once the connection between the I and the Now is severed, the I can freely move into the past or project itself into the future. But this I is not the actual I of immediate presence. It is a Me – an I without the Now. The I and the Now cannot fall out of synchronisation; they always come together. It is self-referring and actual, as it comes into being at the very moment it refers to itself, i.e., Now.

23 By using 'I', we refer to or identify "a thing which is not part of the experience" (Castañeda 2001, 64).

Having that in mind, Descartes' *cogito* argument has a rather implicit presupposition. Indeed, if I think I must exist, because thinking, in this case, is self-referential – I am thinking that I *am* thinking, i.e., I am thinking *now*. Self-referentiality implies that that which is referring exists at the same time as that to which it refers, as they are the same thing. My thinking would not lead to the conclusion about my existence if it were not self-referential, which is precisely what is contained in 'I think'. Lichtenberg's critique (Lichtenberg 2013, 152), that *cogito* should be *cogitatur* is out of place not because of the disagreement if consciousness *per se* is egocentric or egoless, but because it lacks self-referentiality of thinking.²⁴ Without it, it would mean 'thinks therefore exists', which does not follow. There is no thinking (without a self or 'I') which is occurring and will lead to existence. We can here use Lewis' example of two gods with full propositional knowledge of the worlds they are in, but not in an indexical form (Lewis 1979). I argue that *they would not know if they themselves exist*. They would know something along the lines of 'God A exists' and 'God A thinks', but nothing would make them utter "I exist". To that extent, the *cogito* argument in such a world would be false. I exist as long as I think, but I think as long as I can think myself, as I can self-refer, as long as I am *actually* thinking in the present. The validity of evidence is in its very (moment of) execution.

These instances show that there cannot be an actual I "outside" of Now or that is not in the present. Husserl formulates a similar idea in the *nunc stans* character of the I. *Nunc stans* is usually translated as 'standing now' to describe eternity, or what remains unchangeable in the change of time. Husserl uses it to designate the timeless essence of the I. We have pointed out that the pure I must remain identical in every mutability or context. It remains unchangeable in the streaming of life, as always present. The I has temporalisation as its form by which it becomes an enduring, self-constituting, concrete I.

The same ego, now actually present, is in a sense, in every past that belongs to it, another – i.e., as that which was and thus is not now – and yet, in the continuity of its time it is one and the same, which is and was and has its future before it. The ego which is present now, thus temporalized, has contact with its past ego, even though the latter is precisely no longer present: it can have a dialogue with it and criticize it, as it can others. (Husserl 1970, 172)

Again, we come across a moment of alterity in one's consciousness. The content of 'here' and 'now' can change, but of 'I' cannot. I can now say, "On Friday, I was in my office, but I spent my Saturday at home". Empirically speaking, I am (currently) a person enriched by those two memories or experiences. Still, at the same time, on both of those occasions, I was an I for myself, the pure subject identical to the I speaking at this moment (or, in the example of amnesiac, the one who woke up in a hospital). I have both increased my experience (my history, memory, personality, experience,

24 Compare (Zöller 1992).

etc.) and remained the same identical I that is now remembering those experiences. The I is always “now actually present”, as it was on Friday and Saturday. It is temporalising, i.e., being “inserted” into time, having its past and future, and simultaneously persists as the same, identical and immutable I. “To change the conviction is to change ‘oneself.’ But throughout change and unchange the Ego remains identically the same precisely as pole.” (Husserl 2000, 324). It remains numerically identical (Husserl 1977, 161). All of my history is present for me, in the Now, and all of my future opens up for me from this moment onwards. Therefore, the *living present* manifests itself as permanently flowing and streaming and, on the other hand, has a character of or remains *nunc stans* (Kockelmans 1977).²⁵ It is the original, actual I of immediate presence, which is “flowingly-statically present” (paradoxically, permanently present and continuously flowing) and is “enduring through ‘its’ pasts” (Husserl 1970, 185). By being structurally linked with the Now, the I is always actual; it is atemporal yet temporalised.

The I of the past is just a variational mode of the I “as existing”, “as having passed”, i.e., the Me. The I divorced from the Now cannot be an I; rather, it is the Me. ‘I am’ means exactly the same thing as ‘I am now’. In fact, ‘I am’ is the abridged version of ‘I am now’.²⁶ The I experiences itself in the living present, and in this way, it grasps itself, including retention, the now moment and the anticipation. Everything is an object for the I, as the known, for-me or for-I, more precisely, *for-I-now*. It is not the ‘now’ moment of something

25 For Husserl’s notion of *lebendige Gegenwart* see (Olbromski 2018).

26 There is a curious medical case of Clive Wearing, a man who, after sustaining brain damage, lost the ability to form long-term memories (Sacks 2007). His anterograde amnesia hinders him from having memories longer than a few seconds. In an effort to grasp his situation, his last conscious moment, he keeps a journal where he often writes of having just woken up, as if being in a coma for years. He looks at his watch, writes down the time “5 pm. I am awake now” or “I am *now*”, and in the next instance, a few minutes later, that date is crossed out and replaced by a new time and “I am now *perfectly awake*”, “I am conscious”, commenting, “First diary entry I’ve made consciously”, *in infinitum*. This case shows his ability to refer to the present moment, being ‘now’, which for him is the moment of his self-consciousness, his self-reference that he himself is now. Each subsequent crossing of the previous entry is caused by his lack of memory of it, and he is living entirely in the present moment. Each ‘now moment’ is replaced by another, actual or current ‘now moment’. Lines are successions of now’s, the awakenings of his self-consciousness. For him, everything is for the first time, even himself. He does not exist in his past, as he is unable to temporalise himself, to have more than just one moment at a time. His ‘I’ is stuck in the immediate present, in the now. It is a Sisyphus’ work, anguish in trying to grasp the flow of actual ‘now’s’ and the stream of consciousness. In one sense, this case could be compared with Wittgenstein’s idea that “the eternal life belongs to those who live in the present”, for our life has no end in the same way as visual field has no limits (Wittgenstein 2001, 87). In another, Sartre, as if anticipating Wearing’s case, argued that consciousness, as impersonal spontaneity, at “every instant of our conscious lives reveals to us a creation *ex nihilo*. Not a new *arrangement* but a new existence. There is something that provokes anguish for each of us in thus grasping, as it occurs, this tireless creation of existence of which we are not the creators.” (Sartre 2004, 27).

because nothing empirical has as its property that it is 'now' (same as no place is 'here' in virtue of being that exact place). It means the continuous, atemporal, absolute now-actuality. On a similar track was Husserl's student Landgrebe.

'I am' in everyday speech, namely, in the natural attitude, means I am here (*Ich bin da*). This here, this *da*, is made up of those two moments: here (*hier*) and now (*jetzt*). If then one applies the phenomenological reduction to this I-am-here, taken naturally, one will find the transcendental *cogito* as transcendental I-am-here. [...] being transcendental, it is this here (*da*) that makes up the ordinary here and now as such. (Yamagata 1998, 13)

This 'here' and 'now' no longer indicate a 'here' and 'now' in the world. Heidegger also opens up a similar topic, invoking Humboldt. In 'I here', the locative personal designation must be understood in terms of the existential spatiality of Dasein.

W. v. Humboldt has alluded to certain languages which express the 'I' by 'here,' the 'thou' by 'there,' and the 'he' by 'over there,' thus rendering the personal pronouns by locative adverbs, to put it grammatically. (Heidegger 2010, 116)

Locative adverbs "have a relation to the I qua Dasein"; they are primarily existential, not categorical, as "determinations of Dasein". Meaning that they are not primarily locative adverbs or personal pronouns; rather, their "significance is prior to the distinction of locative adverbs and personal pronouns." 'Here,' 'over there' and 'there' are not pure locative designations, yet "characteristics of the primordial spatiality of Dasein". If we interpret Dasein without theoretical distortions, we will see it "immediately in its spatial 'being-together-with' the world". Heidegger also talks about *nearness*, *farness*, *distancing*, *directionality*, and others, all in relation to Dasein. (Heidegger 2010, 102ff). Thus, the question of whether the primordial meaning of locative expressions is adverbial or pronominal is eliminated. In this analysis, 'I' and 'here' are understood as having a deeper foundation, in some sense, as *presence*. There is no immediate differentiation between 'I', *me being here*, and 'here', *where I am*. Again, 'here' is not a property of place; in Heidegger's words, "The here does not mean the where of something objectively present, but the where of de-distancing being with ... together with this de-distancing." (Heidegger 2010, 105). In Heidegger's example, we see that the I and Here are amalgamated, and in Husserl's, that I and Now are.

6. Conclusion: other indexicals as derivative from the I

In van Peursen's interpretation of Husserl's idea of the I, it is "the fundamental ultimate point of reference" departing from which a structural declination towards intersubjectivity is possible (Peursen 1959, 38). Which

is compatible with Shoemaker's idea that each person's system of reference has that person as an "anchoring point" (Shoemaker 2001, 93). And, as we have seen, it is the centre of every context and situation. Only 'I' has its referent permanently fixed, formally and contentwise, and not dependent on anything. I cannot mean with 'I' one thing one day and another the next. The referent of 'I' remains continually the same throughout the life of my consciousness. The ultimate point of reference cannot refer back to itself *as* anything else except the ultimate point of reference or *to* anything other than itself. Its self-reference is outside of any context or is contextual only in the sense that in self-referring it is that which self-refers. If in 'perceiving *x*' I refer to myself as the one who is perceiving, or in remembrance as the one who is remembering, then there ought to be something pure in these cases, i.e., that in all of them I refer to myself as their subject. I refer to myself as the one who refers to himself. The very *act of self-reference is the context*; it is what defines 'I' in its entirety.

We could conclude then, as was the paper's motto, that all indexicals are dependent on the I for their content – 'now', 'here', even 'I' in the empirical sense. 'Here' "refers" to the I in a sense that it represents the place where the I is located; it is the Here of the I, and the same holds for the Now, as the moment at which 'now' is being uttered. Therefore, they also get their formal meaning, along with the empirical content, through the I, i.e., deduced from the act of self-consciousness. On the objective side of things, my office becomes 'here' when I enter it and stops being when I leave; however, from the subjective side, I "carry" my 'here' with me, my hereabouts. The difference with 'now' is that we are stuck with it; the I is "flowing" with the 'now', making it actual. They are meaningless and impossible without 'I', for there is no 'here' if there is no I here.²⁷ We cannot define them unless "through" the I, and the Now and Here are intertwined amongst themselves and with the I, the latter being the pure designation of the subject.

All indexicals rest on self-consciousness, and self-consciousness on actuality. This could be compared with Perry's notion that other indexicals can be eliminated in favour of 'I' and 'now', but that conversely is impossible (Perry 2001, 155). As such, 'I' has a referential, ontological and epistemological priority over all other eliminable pronouns and demonstratives (Castañeda 2001); it is also more fundamental than them, and "other sorts of reference are possible only because this sort of self-reference, that involving the use 'as subject' of 'I', is possible" (Shoemaker 2001, 93). There is no self-consciousness outside of the Now; it is always actual. 'I', 'now' and 'here' are thus all results of the same self-referential or self-conscious act, which further implies that self-consciousness, first-person perspective, and essential indexical cannot be divorced notions. In McGinn's words, "all the indexicals are linked with *I*",

27 There would be no 'now' and 'here', and they would amount to nothing because there would be no self-referential system "pinpointing" itself in time and space.

and 'I' comprises the unique perspective a person has on himself – “to think of something indexically is to think of it in relation to *me*, as I am presented to myself in self-consciousness” (McGinn 1983, 17). And also, if 'I' does not change its content in all of these contexts, those relations are not simply indexical but self-referential. They are always about the same referent, the subject or a thinker, with the world. According to Millikan and Prosser, there is no indexicality in self-referential mental states, and that the error is a result of conflation of egocentricity with indexicality. Smith's mental state represents Smith standing in relation to the object, and there is nothing indexical there (Prosser 2015, 228). I do not fully agree, for as we have seen in the pure use, and in the example of the amnesiac, Smith's mental state does not have the content 'Smith stands in *R* to *o*', but 'I stand in *R* to *o*'. Prosser continues that “their mental states represent each of them as being an *x* such that *x* stands in *R* to *o*”. However, if that is the case, if both Smith's and Jones's mental state represent them to themselves as *x*, then we can replace *x* with 'I' and be back at the start. Or, we can say that Smith's relation to Smith, expressed with 'I' is identical with Jones's relation to Jones, also expressed with 'I', which in turn brings us to the aforementioned pure I.

Furthermore, one could argue that there is only one “real” pronoun and indexical, 'I', all others being derivative and contextual, and having their meaning in relation to the I, by being grounded in self-consciousness – 'you', as posited over and against the I, and together constitutive of 'we' (by saying 'we, I gather others around me'); 'he' being opposed but not necessarily present to the I or direct addressee, etc. However, this requires further investigations into the phenomenology of intersubjectivity, i.e., the realisation that another is the Other, *alter ego*, different yet at the same time identical to me. I can address both Peter and Mark with 'you'; however, my use of 'you' implies that they are both 'I' to themselves. It is not simply the case of an indexical, meaning one person in one context and another in the next, being defined solely by the context; but my use of 'you' presupposes something, the lack of which will not motivate me to utilise it in addressing an animal or inanimate object.

What I aimed to outline in this paper is the idea that the I, Now, and Here and their egocentric counterparts in language, 'I', 'now', and 'here', are structurally intertwined in self-consciousness. They are constitutive parts of the first-person perspective of the subject that uses egocentric indexicals and, as such, are all but impossible without a self-conscious subject, which is, in turn, impossible if not being able to purely self-refer. All paths lead to the actual I of immediate presence, flowingly-statically present, that is constantly purely self-referring, in which it finds its 'now' and 'here'. Meaningfully using the indexical 'I' does not mean correctly referring to oneself, but to *be* self-conscious; correctly referring already presupposes self-consciousness. Connecting phenomenology with indexicality was intended to give a grounding to the essentiality of the indexical in self-consciousness. This would mean that the essentiality is nothing other than (pure) self-consciousness and the foundation for other indexicals.

References

- Brook, Andrew. 2001. „Kant, self-awareness and self-reference.“ In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 9–30. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Castañeda, Hector-Neri. 1981. „Attributions of Self-Reference.“ In *Zeichenkonstitution. Akten des 2. Semiotischen Kolloquiums Regensburg 1978*, by Annemarie Lange-Seidl, 62–66. Berlin: De Gruyter.
- Castañeda, Hector-Neri. 2001. “He’: A study in the logic of self-consciousness.” In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 51–79. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cavallaro, Marco. 2020. “Ego-Splitting and the Transcendental Subject. Kant’s Original Insight and Husserl’s Reappraisal.” In *The Subject(s) of Phenomenology*, by Iulian Apostolescu, 107–133. Cham: Springer.
- Evans, Gareth. 2001. “Self-identification.” In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 95–141. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Frank, Manfred. 1999. „Self-Awareness and Self-Knowledge. Mental Familiarity and Epistemic Self-Ascription.“ In *Avant-Garde Critical Studies, Volume 12: Subjectivity*, by Willem van Reijen and Willem G. Weststeijn, 193–215. Brill.
- Heidegger, Martin. 2010. *Being and Time*. Translated by Joan Stambaugh. Albany: State University of New York Press.
- Husserl, Edmund. 2019. *First Philosophy: Lectures 1923/24 and Related Texts from the Manuscripts (1920–1925)*. Translated by Sebastian Luft and Thane M. Naberhaus. Dordrecht: Springer.
- . 1983. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book: General Introduction to a Pure Phenomenology*. Translated by F. Kersten. The Hague: Martinus Nijhoff.
- . 2000. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. Second Book: Studies in the Phenomenology of Constitution*. Translated by Richard Rojcewicz and André Schuwer. Dordrecht: Kluwer Academic Publishers.
- . 1977. *Phenomenological Psychology. Lectures, Summer Semester, 1925*. Translated by John Scanlon. The Hague: Martinus Nijhoff.
- . 1970. *The Crisis of European Sciences and Transcendental Phenomenology*. Translated by David Carr. Evanston: Northwestern University Press.
- Kant, Immanuel. 1998. *Critique of Pure Reason*. Translated by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press.

- Kaplan, David. 1989. „Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals.“ In *Themes from Kaplan*, by Joseph Almog, John Perry and Howard Wettstein, 481–563. New York: Oxford University Press.
- Kockelmans, Joseph J. 1977. “Husserl and Kant on the Pure Ego.” In *Husserl: Expositions and Appraisals*, by Frederick A. Elliston and Peter McCormick, 269–285. Notre Dame, Indiana: University of Notre Dame Press.
- Lewis, David. 1979. “Attitudes De Dicto and De Se.” *The Philosophical Review* 88 (4): 513–543.
- Lichtenberg, Georg Christoph. 2013. *Philosophical Writings*. Translated by Steven Tester. Albany: State University of New York Press.
- McGinn, Colin. 1983. *The Subjective View: Secondary Qualities and Indexical Thoughts*. New York: Oxford University Press.
- Mead, George H. 1972. *Mind, Self, and Society. From the Standpoint of a Social Behaviorist*. Chicago: The University of Chicago Press.
- Mead, George H. 1913. „The Social Self.“ *The Journal of Philosophy, Psychology and Scientific Methods* 10 (14): 374–380.
- Millikan, Ruth Garrett. 2001. „The myth of mental indexicals.“ In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 163–177. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Olbromski, Cezary Józef. 2018. “Temporalizations of Time: Edmund Husserl’s “Now (ness)”, “Present (ness)”, and lebendige Gegenwart.” *Proceedings of the XXIII World Congress of Philosophy*. 85–89.
- Perry, John. 1977. „Frege on Demonstratives.“ *The Philosophical Review* 86 (4): 474–497.
- Perry, John. 2001. „The problem of the essential indexical.“ In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 143–159. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Peursen, Cornelis van. 1959. „Some Remarks on the Ego in the Phenomenology of Husserl.“ In *For Roman Ingarden: Nine Essays in Phenomenology*, by Anna-Teresa Tymieniecka, 29–41. The Hague: Martinus Nijhoff.
- Prosser, Simon. 2015. “Why Are Indexicals Essential?” *Proceedings of the Aristotelian Society* 115 (3): 211–233.
- Sacks, Oliver. 2007. *The Abyss*. 17 June. <https://www.newyorker.com/magazine/2007/09/24/the-abyss>.
- Sartre, Jean-Paul. 2004. *The Transcendence of the Ego: A Sketch for a Phenomenological Description*. Translated by Andrew Brown. London; New York: Routledge.

- Shoemaker, Sydney S. . 2001. "Self-reference and self-awareness." In *Self-reference and Self-awareness*, by Andrew Brook and Richard C. DeVidi, 81–93. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Taguchi, Shigeru. 2018. „Non-contextual Self: Husserl and Nishida on the Primal Mode of the Self.“ In *The Realizations of the Self*, by Andrea Altobrando, Takuya Niikawa and Richard Stone, 31–46. Cham: Springer.
- Wittgenstein, Ludwig. 2001. *Tractatus Logico-Philosophicus*. Translated by D. F. Pears and B. F. McGuinness. London and New York: Routledge.
- Yamagata, Yorihiro. 1998. „The Self or the Cogito in Kinaesthesia.“ In *Self-Awareness, Temporality, and Alterity*, by Dan Zahavi, 9–19. Dordrecht: Springer Dordrecht.
- Zahavi, Dan. 2021. „From No Ego to Pure Ego to Personal Ego.“ In *The Husserlian Mind*, by Hanne Jacobs, 269–279. London: Routledge.
- . 2014. *Self and Other*. Oxford: Oxford University Press.
- . 2015. „Vindicating Husserl’s Primal I.“ In *Phenomenology in a New Key: Between Analysis and History*, by Nicolas de Warren and Jeffrey Bloechl, 1–14. Cham: Springer.
- Zöller, Günter. 1992. "Lichtenberg and Kant on the Subject of Thinking." *Journal of the History of Philosophy* 30 (3): 417–441.



COGNITIVE DYNAMICS: RED QUEEN SEMANTICS VERSUS THE STORY OF O

Abstract: *It appears that indexicals must have fine-grained senses for us to explain things involving human action and emotions, and we typically identify these different senses with different modes of expression. On the other hand, we also express the very same thought in very different ways. The first problem is the problem of cognitive significance. The second problem is what Branquinho (1999) has called the problem of cognitive dynamics. The question is how we can solve both of those problems at the same time. Vojislav Božičković (2021) offers one solution in which the cognitive dynamics runs through the objects of the attitudes. I discuss this solution and offer an alternative in which the theory of cognitive dynamics has no use for the objects of the attitudes to unify expressions of attitudes. When we say or believe “the same thing” using different modes of expressions, it is by virtue of our deploying a dynamic theory of attitude expression. Like Lewis Carroll’s Red Queen, we must run to stay in place.*

Keywords: *Indexicals, cognitive significance, cognitive dynamics, tracking, red queen semantics, attitudes.*

The topic of indexicals has exercised analytic philosophers for over a century, in part because Frege framed the problems so cleanly. The central issue is this: On the one hand it appears that indexicals must have fine-grained senses for us to explain things involving human action and emotions, and we typically identify these different senses with different modes of expression. On the other hand, we also express the very same thought in very different ways. The first problem is the problem of *cognitive significance*. The second problem is what Branquinho (1999) has called the problem of *cognitive dynamics*. The question is how we can solve both of those problems at the same time.

To see why solving both problems at once is difficult, let’s start by reviewing the problem of cognitive significance. John Perry (1979) has catalogued a number of examples where a theory of cognitive significance is called for, including cases of personal identity, spatial separation, and temporal separation. To illustrate the personal identity case, Perry draws on

an incident described by Ernst Mach, in which Mach once boarded a bus and saw a shabby-looking man, which prompted him to ask himself, “who is that shabby pedagogue,” not realizing he was looking in a mirror. As Perry correctly notes, there is a thought that Mach expresses with an utterance ‘I am a shabby pedagogue’ which he does not express with an utterance of ‘he is a shabby pedagogue.’

The temporal case works in the same way. Developing an example from A.N. Prior (1959), knowing that I have an appointment with the dentist for a root canal today is quite different from knowing that I have an appointment with the dentist on a particular date – say June 1 – if I have lost track of the date and don’t realize that today is June 1. Knowing that the appointment is *today* explains my anxiety as well as why I get up and go to the dentist when I do. Similarly, as Prior pointed out, it is only by knowing the root canal is in *your* past that you can feel relief about it. It is not enough to know that the root canal is earlier than a particular time.

The spatial case also works in the same way. Knowing that a ticking timebomb is in such and such spatial coordinates may not get me up and running in the same way as the knowledge I express when I utter ‘the bomb is here.’

In all these cases, there is pull to introduce fine-grained sense content to explain my difference in attitude in these cases. We can say that we are providing the fine-grained sense content to account for the cognitive significance of our thoughts.

On the other hand, there are cases where the fine-grained indexical content we introduce to account for cognitive significance leads to puzzles. Consider the case, discussed by Frege, where there is a thought that I express with the words ‘Today is a fine day’ and then express that very same thought the following day with the words ‘Yesterday was a fine day’. We want to say that we are expressing the same thought in these two cases, but are we not deploying different senses by using the words ‘today’ and ‘yesterday’? How can we be expressing the same thought? This, again, is the problem of cognitive dynamics.

Engaged individually, the problems of cognitive significance and the problem of cognitive dynamics are not so hard. The real difficulty comes when we want to solve *both* problems, because the apparent solutions seem to be at cross-purposes. The problem of cognitive significance leads us to posit finer-grained sense content to account for the explanation of our actions and emotions. Meanwhile the problem of cognitive dynamics calls us to show how two episodes of thought can have the same fine-grained sense content even though the thoughts are expressed in different ways at different times and places. The second problem calls us to collapse sense contents, or more accurately, it calls us to explain how the very same sense content can be expressed in different ways.

One natural solution to the puzzle is to surrender on either the problem of cognitive significance or the problem of cognitive dynamics. Let's lump these solutions together under the umbrella descriptor of *the surrender strategy*. What is the surrender strategy? One version of the surrender strategy is to say that semantics shouldn't be interested in the business of cognitive significance. In the words of Wettstein (1986) when semantics gets into the business of cognitive significance, it is "resting on a mistake." Another surrender strategy (associated with Heck, 2002) is to give up on the cognitive dynamics part. You actually aren't thinking the same thought in these cases (e.g., when on day one you have a thought which you express by saying 'today is a fine day' and on day two you have a thought which you express with 'yesterday was a fine day'). Maybe there are independent reasons to think that semantics has rested on a mistake, and maybe there are independent reasons to believe that we aren't thinking the same thought from different perspectival positions. My portfolio here is limited. What I aim to argue is that cognitive significance and cognitive dynamics are not in conflict. You can have your cake and eat it too. A compatibilist solution is possible, but we have to get the cognitive dynamics part right.

I'm not the only one who thinks this. In a recent book, Vojislav Božičković (2021) also develops a compatibilist strategy. He and I are fundamentally on the same team here – we both reject the surrender strategy – however, I do not think his account of cognitive dynamics works. His answer comes to this: To give an account of cognitive dynamics, we need to posit some particular thing (I will call it a *cognitive object*), distinct from the thought itself, that can anchor our expression of the same thought at different times and places. I, on the other hand, want to say there can be no such thing, but that we nevertheless do express the same thought in different ways.

To be clear, my target is not really Božičković's proposal so much as a hypothetical class of solutions that attempt to reify something that is going to serve as the hook or binder that keeps the expressions of sense content linked together so that we can say they express the same thing. My solution, by way of alternative, is going to be that while no individual objects can link together expressions of sense content, a theory of cognitive dynamics can explain and predict when these expressions are linked together.

In a bit I'm going to walk through the Božičković proposal in some detail in order to argue that his appeal to (cognitive) objects isn't going to work, and after that I am going to argue that these same objections are going to infect any like-minded theory. Anyone attempting to do what Božičković is attempting to do will run into the same problems.

Before we get into the Božičković strategy, however, I should probably say a word or two about my own strategy here as developed in my (2019) book *Interperspectival Content*. As the title of the book suggests, we are interested in expressing the same thing across diverse perspectival positions. The easy

way to do this, of course, is to dispense with sense content and just stick with objectual or referential or official content or whatever you choose to call it. That is, of course, one of the surrender strategies. Unless supplemented somehow, it surrenders sense content and thus our ability to explain actions and emotions and, as I argued in my book, many other things including ethical decisions, scientific practice, and even information and computation. (I will touch on the latter two cases in a bit.)

So how does interspectival content work? As I said, on my view there are no individual objects that can be the magic content that glues together diverse ways of expressing the same thing. That is, if yesterday I utter ‘today is a fine day’ and today I utter ‘yesterday was a fine day’ I am saying the same thing, but there is no separate object that is unifying those two expressions. What unifies the expressions so that they express the same thing is a great big theory about how we glue these and other expressions together. Let’s call this great big theory *Theta*.

What is *Theta* and how does it work? Let’s illustrate by first talking about belief attributions. Many years ago (1993), Richard Larson and I published a paper called “Interpreted Logical Forms,” in which we covered related ground. On the one hand we want fine grained sense contents to discriminate beliefs (for example the belief expressed by ‘Harvard is a fine school’ from the belief expressed by ‘[Hahvahd] is a fine school’). Larson and I argued that the objects of belief would have to incorporate syntactic information broadly understood to include phonological information as well. But then we noted that there are also cases where we use different expressions to say the very same thing. This is because we want to be able to say that the same belief can be expressed in different ways and indeed in different languages.

I hope it is clear that this is basically the same problem as the problem we encounter with the tension between giving an account of cognitive significance and an account of cognitive dynamics. And it is not unheard of to despair of a possible solution here. In fact, in my very first meeting with Noam Chomsky (in 1983), he argued that any account of the attitudes would founder on exactly this point, and he referenced a paper by Israel Scheffler (1955) which I later learned Chomsky had been citing since he was a post doc, when Scheffler’s paper was still in draft form. Admittedly, it is a hard problem and some of the smartest people in the world think there is no solution to the problem.

These very smart people are not wrong if we take their targets to be the idea that there is some simple solution to the dueling demands of cognitive significance and cognitive dynamics. In a bit, I’ll explain why. But Larson and I weren’t offering up some simple, easy story involving fixed objects of belief. We were offering up a theory of how we dynamically coordinate expressions of meaning without the help of such fixed objects.

The core idea of our theory was that when we engage in belief attribution, we are not denoting a fixed structure that is represented in the mind/brain. We are rather providing information to our communicative partners that will help them construct a theory of the cognitive makeup of the person we are attributing the belief to. This involves lots of background assumptions that we have with our communicative partners about our shared goals in belief attribution and quite a bit of common-sense reasoning, but critically it also involves a strategy we have for constructing mental models of the person we are reporting on and how we tacitly negotiate the expressions we will use to speak of the different components of those models.

What does that mean? Sometimes we are talking about someone that doesn't know the Morning Star is the Evening Star, and we may choose to assign the expressions 'Morning Star' and 'Evening Star' different roles when we are in the business of attributing beliefs to such a person. Why would we do this? Well presumably we would do this because we are interested in the actions and emotions of the person we are reporting on. If we say to them "the Evening Star is a lovely hue of blue these days," we might be interested in whether the person will take out a telescope and look now (this morning) or wait until evening. We are in the business of theory construction, after all, and what we want our theory to do is help our communicative partners better understand the actions and emotions of the person to whom we attribute beliefs.

We don't need to go into a lot of detail for current purposes, but one element of the theory requires comment. It is at this point in the belief attribution process that the issue of the dynamic lexicon enters the story. This is the story that I developed in Ludlow (2014) and the keynote idea of that is that we often engage with our communicative partners by constructing "microlanguages" in which we introduce new terms and modulate word meanings as needed for current conversational purposes. Word meanings thus shift all the time, and much of what we are engaged with in discourse – even in our use of declarative sentences – has more to do with the task of modulating word meanings than it does with the task of directly describing what is going on in the world.

Now, given that our language is very much in flux, even on a conversation-internal basis, it follows that the way we express certain cognitive states will likewise shift on a conversation-internal basis, and since we are in the business of explaining what someone might feel or do, we are in the business of negotiating the use of fine-grained expressions to help accomplish this. So, the expressions we use must of necessity be in flux. As Gareth Evans (1996) remarked regarding indexicals, "we must run to stay still." Here, Evans is referencing the Red Queen from Lewis Carroll's *Through the Looking Glass* – a character who had to keep running just to stay in place. It is an apt metaphor for my project, which I call *Red Queen Semantics*. The idea is that

semantic theory can't deliver up static objects as meanings but must provide a theory of how we use diverse expressions at diverse times and places to say the same thing.

This means that things are complicated. Sometimes we use different expressions to express different things. Sometimes we use the same expression to express different things. Sometimes we use the same expression to express the same thing. And sometimes we use different expressions to express the same thing. It all depends on the theory we are currently constructing, with our conversational partners, to explain the actions and emotions of the agent we are talking about (sometimes, of course, talking about ourselves).

Does this mean that we have dispensed with beliefs and thoughts and meanings? Not at all; it simply means that beliefs can't be latched onto permanently with a single use of language. It is a dynamic process. There is no one expression you can use to say the same thing in different places at different times. There is no "view from nowhere" in which you can permanently articulate belief contents or thought contents or meanings.

In Ludlow (2019) I took this basic idea about belief attribution and observed that it is basically already designed to account for our use of indexicals in explaining the actions and emotions of other agents (and ourselves, for that matter). You simply have to understand that things like essential indexicals are handy tools that we have for constructing these dynamic theories. From this perspective, essential indexicals are not weird outliers in our semantic theory. They are core cases, and indeed are completely normal and native elements to a Red Queen semantic theory.

The big point in all this, the one that needs to be reiterated, is the idea that there can be no objects that serve as perspective-neutral descriptors of the contents of indexical thoughts. Indexicals (I would prefer to call them perspectival expressions) are just components in theories that we are constructing on the fly (similar to Davidson's notion of passing theories). The theory will have to be packaged in different ways at different times and in different places and for different persons. This means that semantic theory is not the business of constructing aperspectival sets of rules for fragments of natural language, but rather is the business of understanding the mechanisms by which we are able to construct these very perspectival passing theories.¹

One might wonder if it really must be so complicated. Isn't there some other way to thread this needle and account for cognitive significance and cognitive dynamics without all the talk about micro-languages and shifting theories that must be expressed in different ways at different times and different places? This leads us to Božičković (2021).

Božičković has an admittedly simpler theory for unifying the different ways we have of expressing the same thought or sense content. I want to

1 This doesn't mean we can't have fragments. It means that fragments must be continually updated, and that semantic *theory* must therefore be a theory of how we accomplish this.

spend a fair bit of time on this proposal because I believe it provides the best possible effort to give such a simpler account, and thus helps us to see what the difficulties here actually are.

Božičković wants to bind together the different ways of expressing the same thought as I do, but without my big, complicated theory Theta and the Red Queen Semantics. He has a different story of how expressions of sense are bound together. What is his binder? The idea is that we express the same thought about some object *o* from different perspectival positions just in case we unreflectively assume that *o* is held constant across these perspectival positions.

So, for example, what binds together my yesterday's thought, expressed by an utterance of 'today is a fine day' with today's thought, expressed by an utterance of 'yesterday was a fine day', is the day itself. It is the object *o* that is deployed to bind together the different expressions of sense.

It isn't quite this simple, however, because the object alone is not a sufficient condition to bind the sense expressions together. For example, I might come to suspect it is not the same object in the two thought episodes. We will get to cases like this in a bit. But first let's stay focused on *o* itself.

The first issue is the availability of the object *o* and what is required for us to have (presumably) referential thoughts about it. It seems that reference to *o* has to come on the cheap, because we are going to discover cases in which we don't have causal access to *o*. As Božičković (2022) notes, I might have thoughts about tomorrow. For example, I might have a thought which I will express as 'tomorrow will be a fine day'. Tomorrow isn't in my causal history, so how do I refer to it? Božičković opts for a more liberal story about reference from Hawthorne and Manning (2012).

So far, all this is coherent enough. We can spot him a liberal account of reference. But here is where things start to get tricky. Sometimes we have thoughts which aren't linked to any existing objects at all. Consider Peter Geach's (1967) case of Hob, Nob, and Cob, who all believe a witch has been blighting their mare. But there is no witch. So, there is no object to unify their beliefs. To make the issue particularly sharp, let's add a temporal dimension to it. Hob has a thought on day one, which he expresses by uttering 'today the witch is blighting my mare'. On day two he reflects on his earlier belief and expresses it by uttering the following: 'yesterday the witch was blighting my mare'. Presumably, we want to say that Hob had the same thought in these two episodes, but what is the object *o* linking those thoughts together? In response, Božičković offers the following:

I claimed that in order for a belief about an object (or day) to be retained, it is required that the thinker's belief has the relevant internal continuity. Similarly, there is an internal interpersonal continuity between our present interlocutors' beliefs that binds these different expressions into a single sense content (which lacks reference) that accounts for the

‘intentional identity’ that Geach (1967), Edelberg (1986) and others have been concerned with. (2022: 77)

But what provides this interpersonal continuity? What is the thing that unifies the thoughts of Hob, Nob, and Cob, and which ultimately unifies the two episodes of thought by Hob on day 1 and day 2? We can agree that the thoughts must be unified, because we have rejected the surrender strategy and both Božičković and I have maintained that an account of cognitive dynamics will unify everything. And Božičković wanted to introduce an object *o*, that could unify our beliefs in the standard cases (today is a fine day/yesterday was a fine day) but now we get into territory where there is no such object to knit the belief episodes together. It isn’t really helpful to be told that there is “an internal interpersonal continuity.” That is the thing that the account of cognitive dynamics was supposed to explain.²

All of this of course leads to the following question: If you *can* successfully link expressions of thought together without the help of object *o*, why do you need the object at all? You didn’t need it for the Hob, Nob and Cob case, so why do you need it for the today/yesterday case? Perhaps the answer is that one needs some sort of intentional object employed in these cases, or alternatively it is not *o* itself that links things together, but a representation of *o*. This is at least suggested by the passage in Božičković’s reply in which he speaks about mental files. Whether this is the strategy or not, it is certainly an option. One can say that it is not really the object *o* that glues things together, but rather the file that we keep on *o*. If there is no object *o*, you can still have the file. I don’t mean to suggest that using mental files is the only strategy here, but it does seem that there must be some representational layer doing the work of unifying things.

But notice now that this is not only an issue in cases of witches, but it is an issue in cases where the object about which we have beliefs is not only real but even quite salient. Why so? Well, as Božičković notes, the critical issue is whether there is an “internal continuity.” But this can’t come for free, even if the object *o* is there. Let’s start with the following passage from Božičković, in which he considers whether the thought I have about a bottle of wine in my refrigerator is stable, or whether it must come apart (as it would when I incorrectly suppose the bottle has been switched):

Suppose the bottle has remained the same and I unreflectively take it for granted that it is the same bottle from Monday to Tuesday. The sense that I am entertaining is thereby the same throughout. Once I have abandoned this assumption, the senses split and are no longer the same. (2022: 79)

2 The “internal interpersonal continuity” line is also troubling for other reasons. First of all, internal to who or what? How can it be internal and interpersonal at the same time, barring some sort of Hegelian story in which Hob, Nob, and Cob share a mind to which the intentional object is internal?

The part I want to focus on here is the expression ‘take it for granted that it is the same bottle’. Just how innocent is this expression? Or to put it another way, what does he mean by ‘take it for granted’ here? Here is a possible story. Above, in the case of Hob, Nob, and Cob we envisioned that Hob kept a mental file on the supposed witch. Notice that the same sort of thing seems to be required here. I open the fridge on the second day, and I either keep the original file for the bottle or open a new file. Notice that this can’t be the product of some reflective decision. If I merely entertain the idea that it might be a new bottle, I must open a new file.

In a little bit I am going to press this inquiry further and interrogate what it means to open a new file, but notice here that fundamentally we aren’t that interested in the object *o* itself, so much as in the file (or files) we keep on *o*. This is so not merely to help us handle the Hob/Nob/Cob case, but even the simplest cases in which we wonder if a switch has taken place (wine bottles etc.). It might not be files, but it has to be some representational object distinct from the object *o*. Let’s come up with a generic name for whatever this is – $f(o)$ – to either indicate the file on *o*, or some function that maps to a representation of *o*.

What’s wrong with this idea? Well, the general strategy is this: Introduce some object and have it serve as the glue or binder rings that hold together the different ways of expressing the same sense content. But we need to ask, what exactly are we talking about when we talk about $f(o)$? The danger here is that the story of $f(o)$ will collapse into a story that is completely vacuous. To wit, that the story of $f(o)$ is whatever we need to keep the sense expressions unified. It is “that which binds these things together.” It is in that case the semantic equivalent of “dormitive powers.”

Before we take a deeper dive into the tricky business of mental files, I want to dwell just a bit longer on our bottle of wine, for there is a lot going on here, and any account of $f(o)$ will have to deal with these details. Let’s say that at time t_1 I put the bottle of wine in my refrigerator, uttering to myself ‘this wine is supposed to be delicious’. The next day, let’s call this time t_2 , I open the fridge and look at the bottle and again utter ‘this wine is supposed to be delicious’, without wondering if it is the same bottle or if a trickster has moved it or if I have forgotten replacing the wine in the fridge. In that case, an utterance of ‘This wine is supposed to be delicious’ at t_1 and my utterance of ‘This wine is supposed to be delicious’ at t_2 can express the same thought, but just in case it is the same bottle. Recall that as soon as I wonder if the bottle has been switched, the sense contents must come apart. Representations (files) would be multiplied to accommodate this possible expanded ontology.

There are things to puzzle over here. For example, in Ludlow (2022) I worried that this strategy might undermine our past and future thoughts. So, for example, suppose that at t_1 I put a bottle of wine in the fridge and utter ‘that wine is supposed to be delicious’. At t_2 I remind myself what I said and

utter ‘Yesterday I thought that wine is supposed to be delicious.’ I can do that on Božičković’s proposal because it has never occurred to me that there has been a change in bottles. But now suppose that on Friday (let’s call it time t_3) I learn that my roommate has been changing things in my fridge, drinking things in the evenings and replacing them, so that I come to wonder if it was indeed the same bottle. Do the senses of my earlier two thought events now come apart? I argued that they must come apart because I am no longer in a position to express both thoughts in the same way. I have doubts about the stability of the underlying referential contents. That is to say, whatever sense content I express after t_3 with an utterance of ‘I thought that wine is supposed to be delicious’, it cannot be the same as both the t_1 thought and the t_2 thought. On Friday I can express the Monday thought, or I can express the Tuesday thought, but I cannot express both at the same time, because the sense contents have retroactively come apart.

Božičković is unimpressed by this objection:

The fact that [the sense contents] will split in the future does not affect the fact that prior to this I keep thinking of the bottle via the same sense. (To allow the future divisions of sense to affect my current state of mind is also to admit of senses a difference in which is not transparent which is at odds with the transparency thesis that I hold on to in the book.) (2022: 79–80)

But here is the problem. It is all very well to say that my thought episodes are unified at the earlier time and it is also very well to say that my thoughts must be separate at a later time. The problem comes in when we want to say that I can recall my earlier thoughts after the switched-bottle reflection has been induced in me. I believe (and I assume Božičković believes) that I can recall my t_1 and t_2 thoughts at a later time, but how does this work given that I no longer have the correct sense content in which to clothe those thoughts?

Here we need to speculate. Maybe, after using only one object file, call it file $f(o)$, at t_1 and t_2 , and after introducing separate files, call them files $f_1(o)$ and $f_2(o)$, at time t_4 I can still deploy that unified file $f(o)$ and it is this file that I use for my memorial access to the earlier thoughts.

One issue here is that the project we are embarked on is not exactly representationally austere. We not only have to split files as needed, but we also have to keep the original files around in case we ever need to recall our pre-fission thoughts. This would be true in the opposite direction as well. If files are fused you have to keep around the pre-fusion files in case you ever decide to recall an earlier pre-fusion thought (for example, before you knew *The Morning Star* was *The Evening Star*).

A further issue involves the question of how it is that I can access a unified sense content at t_4 , when it was supposed to be automatic for these sense contents to come apart the second I begin to wonder whether there has been an object switch.

But the real issue is the issue I raised earlier, and which has now returned with a vengeance. We were earlier worried about whether we have any independent handle on these files or whatever the representations $f(o)$ are. Now, whatever they are, they seem to be positively magical. Did we actually add something contentful when we called the secret ingredient for cognitive dynamics “ $f(o)$ ” as opposed to simply “the sense content”? I don’t think so. It seems that we only know to introduce files when we think we need to introduce a new sense content to distinguish our thoughts and we know to unify the files again when we need a single sense content.

You might think that shifting to talk of files or representations gets us further because it embeds the talk in a field of empirical enquiry – cognitive science – and we can understand files to be data structures of some form within this field, but this is in fact just a way of repackaging the same problem. There is no independent way to ground the existence of the requisite data structures/files. There is also no aperspectival way to identify the relevant data structures/files. Let’s take the latter issue first.

Let’s assume for the moment that the idea of data structures/files is unproblematic and that we can ground data structures with no problem (we will question this assumption in a bit). We will say that data structures are semi-stable syntactic states of a computational system that encode the information we have about objects in the world. Returning to our example of the bottle of wine, the proposal on the table is that we have a single data structure for the wine-related thoughts at times t_1 and t_2 – this was $f(o)$. Then there are the $f_1(o)$ and $f_2(o)$, which are the files that get deployed after t_3 . The issue is that it isn’t enough for me to *have* these files; crucially I have to deploy the correct file. Let’s say that at t_4 I recall my t_1 thought. So I must deploy the file $f(o)$. But what makes it so that I am applying that file as opposed to one of the others? It seems that the only handle we have on that file is to know that it is the file corresponding to the thought that *I* had at *that* time in *my past*. In other words, I need the sense contents to simply identify which file is $f(o)$.

There is an even deeper problem, this one involving the files themselves as opposed to our previous concern about deploying the correct file. From a non-technical perspective, the problem is this: What makes it true that the data structures/files carry the information they need to here? Does the content of such data structures come for free? It does not.

From the perspective of the foundations of the theory of computation, the information carried by the data structures cannot be grounded in the low-level physical properties of the computational system, but rather must be grounded in terms of the legibility of the inputs and outputs of the computational system, which in turn means we are interested in the legibility of the inputs and outputs *to us*. And this includes the legibility of the perspectival content that may be carried by the computational system. That content isn’t grounded in the physical properties of the system, but rather

in subjective – indeed perspectival – properties that we build the systems to encode. The perspectival properties come first, which is to say that the sense content must come first. You can't ground the sense content of the data structure without having an external anchoring of the the sense content.

This is a thesis I advanced in Ludlow (2019; Chapter 5), and while there isn't space to develop the argument here, I can at least offer the elevator pitch for those that are interested. The basic idea is that indexical content – what I would prefer to call perspectival content or interspectival content – not only figures in our accounts of human action and emotion, but it also figures in our accounts of ethics (for example I must know that something is a rule *for me*), and scientific practice at the experimental level (at the simplest level, I have to know that something is *my* experiment across time), and ultimately it grounds our accounts of information and thus computation as well (because our best theory of information is subjective and perspectival).

Information is subjective and perspectival? Yes! This is a deep point that runs through Shannon's (1948) theory of information,³ and even through Boltzmann's theory of entropy in statistical thermodynamics,⁴ on which Shannon's theory is grounded. As Galistel and King (2009) put the idea, "the information communicated by a signal depends on the receiver's (the subject's) prior knowledge of the possibilities and their probabilities. Thus, the amount of information actually communicated is not an objective property of the signal from which the subject obtained it!" In Ludlow (2009) I supplemented this idea with the observation that the information communicated is not merely a subjective property but a perspectival property as well, for there is also the matter of whether the communicated information is *for me*. That isn't an objective property of the signal either.

Following Galistel and King, we can illustrate the point with the example of the (somewhat apocryphal) story of Paul Revere, from the American Revolution. When Paul Revere sees the lights in the Old North Church he understands that the lights mean that the British are coming by land and by sea, but it is not some objective property of the lights that encodes this. It is only by his prior understanding of possible messages that Revere understands what they mean. But note also that it is not an objective property of the light signals that they carry the information that this is a message *for him*. Nor is it an objective property of the signal that it carries the information that the invasion is happening *now*. This too is something Paul Revere must know to be a possible message.

So far, I have pointed out that the information carried by a data structure is a subjective and perspectival property; it is not an objective property of the computational system. There is nothing about the lights in themselves that tell

3 See also Shannon and Weaver (1949).

4 See Boltzmann's 1886–1889 lectures on gas theory, translated and published in Boltzmann (1964).

you the British are coming. You have to know what the possible messages are. Similarly, given two data structures, $f_1(o)$ and $f_2(o)$, there is a serious question of what those structures represent, and whether they represent two different sense contents is not an objective property of the data structure itself. We have to have a prior independent understanding of the sense contents.

But the problem is even worse than this, because we can also interrogate the nature of the data structure itself. By virtue of what do we have two separate data structures instead of one? Or any data structure at all? The data structures cannot be individuated by the microphysical states of the system. This is a point that is made by Kripke (1982) in his reconstruction of Wittgenstein's rule following argument, and in particular in his point that computational states can't supervene on internal low-level properties of the physical system. This needn't lead us to skepticism, however. It merely means that the very data structures of a computational system depend on the information being communicated through the system, and as with all information, this is a subjective and quite perspectival property turning on either the intentions of the designer of the system and/or whoever reads the input and output of that system.⁵

What this means for current purposes is that you can't ground sense content in a computational account of the mind/brain because the computational or information-theoretic account (necessary for an appeal to data structures and files) is itself anchored by sense content. In other words, data structures can't anchor sense content because we need perspectival sense content to individuate the relevant data structures. Perspectival content comes before information-theoretic content (or at least the information-theoretic content we need here). To think otherwise is to put the cart before the horse. You need an account of perspectival content before you can start talking about data structures and files that might do the work of distinguishing fine-grained thoughts that trade in indexical content.

Now, I want to make it clear that I am not at all sure Božičković would push this line about data structures as hard as I have here. My target is perhaps more correctly understood as someone who might make a herculean effort to find some object to distinguish sense contents and ground them in

5 The following passage from Kripke (1982) develops this idea.

Actual machines can malfunction: through melting wires and slipping gears they may give the wrong answer. How is it determined when a malfunction occurs? By reference to the program of the machine, as intended by the designer, not simply by reference to the machine itself. Depending on the intent of the designer, any particular phenomenon may or may not count as a machine 'malfunction'. A programmer with suitable intentions may even have intended to make use of the fact that wires melt or gears slip, so that a machine that is 'malfunctioning' for me is behaving perfectly for him. Whether a machine ever malfunctions and, if so, when, is not a property of the machine itself as a physical object but is well defined only in terms of its program, as stipulated by its designer. Given the program, once again the physical object is superfluous for the purpose of determining what function is meant. (1982: 34–35)

data structures within cognitive science. But I do want to make the point that there are no strategies that can accomplish what the advocate for such a view needs to accomplish here. The objections I have raised here will apply to any attempt to provide some object, whether *o*, or a file, or some alternative object *p* or *q* or *r* or *s*. Any attempt to execute a general strategy like that of Božičković is going to meet the same end.

There are no cognitive objects that can do the work that Božičković needs done, because we need to lean on sense content to identify them. And if we rely on sense content to identify them (as we did in the case of computational data structures), then we can't use them to ground our account of sense content – we would immediately get trapped in a vicious circle. On the other hand, if we try to rely on objects that are not cloaked in sense contents (e.g. just *o* itself), then of course they will not be fine-grained enough to account for cognitive significance.

I believe, like Božičković, that we *can* give accounts of cognitive dynamics; my point is that the theory of dynamics must do all the hard work itself, without the help of cognitive objects to anchor meaning and sense content. The Red Queen needed to keep running if she wished to stay in place. There were no places she could sit and rest if she wanted to stay in place. Our theory of cognitive dynamics must keep updating modes of expression to say the same thing, and sadly, there are no cognitive objects that can allow the theory to sit and rest if it wants to target a particular thought. Like the Red Queen, the theory must forever stay in motion.

References

- Boltzmann, L. (1964). *Lectures on Gas Theory 1896–1898* (trans. S. Brush), Berkeley: University of California Press.
- Božičković, V. (2021). *The Indexical Point of View*. New York: Routledge.
- Božičković, V. (2022). “Reply to Ludlow.” *Manuscrito* 45 (3), pp.74–83.
- Branquinho, J. (1999). “The Problem of Cognitive Dynamics.” *Grazer Philosophische Studien* 56, pp. 29–56.
- Edelberg, W. (1986). “A New Puzzle about Intentional Identity.” *Journal of Philosophical Logic* 15, pp. 1–25.
- Evans, G. (1996). “Understanding Demonstratives.” In *Collected Papers*. Oxford: Oxford University Press, pp. 291–321.
- Galistel, C.R., & King, A. (2009). *Memory and the Computational Brain: Why Cognitive Science Will Transform Neuroscience*. Chichester: Wiley-Blackwell.
- Geach, P. (1967). “Intentional Identity.” *The Journal of Philosophy* 74, pp. 627–632.
- Hawthorne, J., & Manley, D. (2012). *The Reference Book*. Oxford: Oxford University Press.

- Heck, R. (2002). "Do Demonstratives have Senses?" *Philosophers' Imprint* 2, #2, <www.philosophersimprint.org/002002/>.
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.
- Larson, R., & Ludlow, P. (1993). "Interpreted Logical Forms." *Synthese* 95, pp. 305–355.
- Ludlow, P. (2014). *Living Words: Meaning Underdetermination and the Dynamic Lexicon*. Oxford University Press.
- Ludlow, P. (2019). *Interperspectival Content*. Oxford: Oxford University Press.
- Ludlow, P. (2022). "Reflections on The Indexical Point of View: On Cognitive Significance and Cognitive Dynamics, by Vojislav Božičković." *Manuscripto* 45 (3), pp. 60–73.
- Perry, J. (1979). "The Problem of the Essential Indexical." *Noûs* 13, pp. 3–21.
- Prior, A.N. (1959). "Thank Goodness That's Over." *Philosophy* 34, pp. 12–17.
- Scheffler, I. (1955). "Synonymy and Indirect Discourse." *Philosophy of Science* 22, pp. 39–44.
- Shannon, C., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, IL: University of Illinois Press.
- Shannon, C. (1948). "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27, pp. 623–656.
- Wettstein, H. (1986). "Has Semantics Rested on a Mistake?" *Journal of Philosophy* 83, pp. 185–209.



DE SE THINKING AND MODES OF PRESENTATION¹

Abstract: *De se* thoughts have traditionally been seen to be exceptional in mandating a departure from orthodox theories of attitudes. Against this, skeptics about the *de se* have argued that the *de se* phenomena demand no more of our theories of attitudes than traditional Frege cases. In this camp one view is that the *de se* can be accounted for by MOPs in the same way that MOPs can account for how it can be rational to believe, for instance, "Hesperus is shining" while also believing "Phosphorus is shining." This paper formulates some minimal conditions that *de se* MOPs must have in order to explain the relevant *de se* phenomena. Some potential replies are answered. I conclude that *de se* MOPs are not exceptional.

Keywords: *de se*, modes of presentation, attitudes, indexicals.

1. Introduction

A highly influential view in philosophy holds that there is a special category of *de se* thoughts, also known as *indexical*, *egocentric*, or *first-person* thoughts. Such thoughts are said to be about oneself in a particular sense that makes them substantially different from thoughts one might have about other things, and indeed different from other thoughts one might have about oneself.²

To echo Kaplan's (1989) well-known example, imagine that you see someone reflected in a shop window, and you realize that they are about to be hit by snow falling from the roof. You think, "They're about to be hit!" But what you did not realize was that it was yourself you saw. Had you realized this, you would have thought, "I'm about to be hit!" And even though, in a familiar sense, both thoughts are about you, only the latter is a *de se* thought,

1 I am grateful to audiences at Uppsala University and Lund University for discussion. Thanks, in particular, to Anandi Hatiangadi, Carl Montan, Jessica Pepp, Nils Franzén, and Will Gamester for very useful feedback.

2 Proponents of this view, or close relatives of it, include Castañeda (1966), (1975), Lewis (1979), Perry (1993 [1979]), Evans (1982), McGinn (1983), Kapitan (1995), Velleman (1996), Burge (2004), Ninan (2010), (2013), (2016), Owens (2011), Recanati (2012), Weber (2014).

that is, a thought about you *as yourself*. Moreover, only the latter, *de se* thought will make you move out of the way or cover your head.

Examples of this kind have motivated two claims about *de se* thoughts, which I label as follows:

Essentialism

De se thoughts are necessary for intentional action.

Exceptionalism

De se thoughts mandate a departure from orthodox accounts of attitudes.

Both of these views have been challenged by philosophers who are skeptical about the *de se* as a special category of thoughts.³ This paper concerns Exceptionalism, and I will not have anything to say about Essentialism.⁴

Anti-Exceptionalists, such as Cappelen and Dever (2013) and Magidor (2015), have suggested that the phenomena that motivated Perry (1993 [1979]), Lewis (1979), and others, to endorse versions of Exceptionalism can be explained by the resources that we already need to explain ordinary Frege cases. Since any theory of attitudes must explain (away) Frege cases, Anti-Exceptionalists of this stripe conclude that the *de se* is not exceptional.

Frege cases are standardly handled in terms of so-called *modes of presentation* (MOPs). Accordingly, Anti-Exceptionalists have suggested that the *de se* can be handled by MOPs in the same manner. Roughly, David can rationally believe that he is about to be hit under the MOP "They're about to be hit" while disbelieving that he is about to be hit under "I'm about to be hit." Yet these arguments have been mainly negative. Little or no attention have been given to showing how *de se* phenomena can be accounted for within orthodox views of attitudes.

My aim in this paper is to spell out how first-person or *de se* MOPs need to be understood in order to explain the relevant phenomena. One can see this as a way of delineating an Anti-Exceptionalist position. But more generally, asking what the *de se* demands of a traditional view of attitudes in terms of MOPs is a way of asking what precisely is demonstrated by the *de se* cases that have been appealed to.

Section 2 reviews the opposition between Exceptionalism and Anti-Exceptionalism and introduces the main idea of accounting for the *de se* phenomena in terms of MOPs. In Section 3 I describe three features that *de se* MOPs must have in order to do the work the Anti-Exceptionalist claims they can do. Section 4 considers some potential Exceptionalist replies.

3 De se skeptics include Stalnaker (1999 [1981]), Millikan (1990), Devitt (2013), Douven (2013), Cappelen and Dever (2013), Magidor (2015).

4 For some recent defenses of (different versions of) Essentialism, see e.g. Prosser (2015), Babb (2016), Bermudez (2017), Morgan (2019), Stokke (in press-b).

2. Exceptionalism and Anti-Exceptionalism

2.1. The Doctrine of Propositions

Both Perry (1993 [1979]) and Lewis (1979) explicitly took the *de se* phenomena they identified to motivate Exceptionalism. That is, they both thought that *de se* thinking requires an explanation that goes beyond orthodox theories of attitudes. Of course, this claim begs the question of what counts as orthodox theories of attitudes.

Perry held that "the essential indexical is a problem for the doctrine of propositions." (Perry, 1993 [1979], 37) Following Ninan (2016), we can spell out what Perry meant by "the doctrine of propositions" as the conjunction of three claims:⁵

The Doctrine of Propositions

Two-Place

Attitudes are two-place relations between subjects and contents.

Frege's Constraint

If a subject can rationally have a belief she could express by "S" without having a belief she could express by "S'", the two beliefs have different contents.

Absoluteness

Contents vary in truth value only with worlds.

The Doctrine of Propositions is consistent with different theories of attitudes.

One view that obeys The Doctrine of Propositions takes a content to be a pair of an absolute proposition and a MOP, or what Salmon (1986) called a "guise," and takes attitudes to be two-place relations between subjects and such pairs. Further, MOPs are seen as truth-conditionally inert.⁶ I label this view *Minimal Propositionalism*:

Minimal Propositionalism

- (i) A content is a pair $\langle p, *m^* \rangle$ of an absolute proposition p and a MOP $*m^*$.
- (ii) Attitudes are two-place relations between subjects and such pairs.
- (iii) MOPs are truth-conditionally inert.

For example, suppose Sue believes "Hesperus is a planet" but does not believe "Phosphorus is a planet".⁷ An adherent of Minimal Propositionalism analyzes this situation as in (1a-c) (where v is Venus).

5 This arguably deviates slightly from Perry (1993 [1979], 36), yet I take this to be a clearer and more useful way of spelling out the traditional view that Perry was targeting than the one Perry originally gave himself.

6 Needless to say, MOPs do play a role in the truth conditions of attitude ascriptions. Yet the truth conditions of the attitudes themselves are not affected by MOPs, on this view.

7 "I use "..." to indicate how a subject would express a particular belief, or what she would assent to, or how the thought might appear in her inner speech. It is important to

- (1) a. Sue believes $\langle v$ is a planet, \ast Hesperus is a planet \ast \rangle .
 b. Sue does not believe $\langle v$ is a planet, \ast Phosphorus is a planet \ast \rangle .
 c. For any w , $\langle v$ is a planet, \ast Hesperus is a planet \ast \rangle is true at w iff $\langle v$ is a planet, \ast Phosphorus is a planet \ast \rangle is true at w .

This represents the general idea that while the belief "Hesperus is a planet" is about Venus, in the sense that it is true if and only if Venus is a planet, it essentially involves a way of thinking about Venus, roughly understood as the role Venus plays in one's cognitive life. As such, the belief "Phosphorus is a planet" is equally about Venus, and has the same truth conditions, but involves a different way of thinking about Venus.

Minimal Propositionalism satisfies The Doctrine of Propositions. But moreover, it is consistent with The Doctrine of Propositions to hold different theories of the propositional component of the pairs that are identified as the contents of attitudes. One version takes the proposition $\langle v$ is a planet \rangle to be a structured entity that includes Venus, the planet itself. Following Recanati (1993), (2012) this view might be called "Neo-Russellianism."⁸ Yet one can also accept The Doctrine of Propositions and hold that $\langle v$ is a planet \rangle is a set of worlds.⁹

Another kind of view that satisfies Perry's doctrine of proposition instead holds that thinking "Hesperus is a planet" involves a different proposition than thinking "Phosphorus is a planet," even though these propositions have the same truth-conditions across worlds. Recanati (1993), (2012) calls such views "Neo-Fregean." So the Neo-Fregean analyzes Sue's situation as in (2a–c).

- (2) a. Sue believes *Hesperus is a planet*.
 b. Sue does not believe *Phosphorus is a planet*.
 c. For any w , *Hesperus is a planet* is true at w iff *Phosphorus is a planet* is true at w .

This view also satisfies The Doctrine of Propositions.

Evaluating to what extent these views are substantially different, or their respective merits, will not concern us here.¹⁰ In what follows I will focus on Minimal Propositionalism.

distinguish beliefs in this sense from MOPs, since, for instance, one can rationally believe "Paderewski, the pianist, has musical talent but Paderewski, the statesman, doesn't," if one does not realize that the pianist and the statesman are one and the same. In such a case one associates with the name two distinct MOPs, such as \ast Paderewski₁ \ast and \ast Paderewski₂ \ast .

8 Cappelen and Dever (2013, 70) call this view "Naïve Russellianism."

9 Cappelen and Dever (2013, esp. ch. 5) seem to have this view of propositional contents in mind.

10 See Recanati (1993), (2012) for thorough discussion. And see also Cappelen and Dever (2013, §4.4).

2.2. Frege Cases and Perry Cases

Exceptionalism has traditionally been motivated by cases. Here is one of Perry's classic examples:

Messy Shopper

I once followed a trail of sugar on a supermarket floor, pushing my cart down the aisle on one side of a tall counter and back the aisle on the other, seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch. [...] I believed at the outset that the shopper with a torn sack was making a mess. And I was right. But I didn't believe that I was making a mess. (Perry, 1993 [1979], 3)

In order to argue that "there is nothing deeply central about indexicals here" Cappelen and Dever compare the Messy Shopper case with the following story:

Messy Superman

Pushing my cart down the aisle I was looking for CK to tell him he was making a mess. I kept passing by Superman, but couldn't find CK. Finally, I realized, Superman was CK. I believed at the outset that CK was making a mess. And I was right. But I didn't believe that Superman was making a mess. (Cappelen and Dever, 2013, 33)

Messy Superman is a Frege case. The subject believes "Clark Kent is making a mess" but does not believe "Superman is making a mess" without irrationality, even though "Clark Kent" and "Superman" co-refer. In turn, as I shall say, Messy Shopper is a *Perry case*. That is, a case in which a subject believes "... I..." but disbelieves "... n...", where *n* is a referential term, without irrationality, even though "I" and "n" co-refer, given the context.

Commenting on Messy Superman, Cappelen and Dever write,

the ease with which Frege counterparts [of the de se examples] can be generated makes at least a prima facie case that the Perry/Lewis/Prior-style cases simply are familiar substitution puzzles and that nothing new is brought out that distinguishes indexicals from other referring expressions with respect to opacity. (Cappelen and Dever, 2013, 68)

Correspondingly, Magidor (2015) writes,

One could plausibly argue that any account of ordinary propositional attitudes would need to involve modes of presentation (or similar devices), due to Frege's puzzle. One could then concede that there is one (or one kind of) mode that is first-personal, and that this mode is particularly important for attitudes that play a role in intentional

action. However, this in itself does not require any revision of our standard account of attitudes or attitude ascriptions. (Magidor, 2015, 258)

In other words, the Anti-Exceptionalist claims that Frege cases and Perry cases are on a par with respect to what they demand of theories of attitudes. More particularly, I take Anti-Exceptionalism to be the view that Minimal Propositionalism suffices for explaining Perry Cases.

2.3. Frege's Constraint and Modes of Presentation

The kind of opacity demonstrated by Frege cases traditionally motivated Frege's Constraint, and MOPs are a standard way of reacting while preserving Two-Place and Absoluteness. Accordingly, the Anti-Exceptionalist will analyze Messy Superman as in (3).

- (3) a. a believes $\langle s$ is making a mess, \ast Clark Kent is making a mess \ast \rangle
 b. a does not believe $\langle s$ is making a mess, \ast Superman is making a mess \ast \rangle
 c. For any w , $\langle s$ is making a mess, \ast Clark Kent is making a mess \ast \rangle is true at w iff $\langle s$ is making a mess, \ast Superman is making a mess \ast \rangle is true at w .

So if the Anti-Exceptionalist is right that Messy Shopper does not demand more of our theories of attitudes, we should be able to give a parallel analysis, as in (4).

- (4) a. j believes $\langle j$ is making a mess, \ast The shopper with the torn sack is making a mess \ast \rangle
 b. j does not believe $\langle j$ is making a mess, \ast I'm making a mess \ast \rangle
 c. For any w , $\langle j$ is making a mess, \ast The shopper with the torn sack is making a mess \ast \rangle is true at w iff $\langle j$ is making a mess, \ast I'm making a mess \ast \rangle is true at w .

Similarly, take Lewis's example of the two gods:

Two Gods

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts. (Lewis, 1979, 139)

Lewis saw this case as motivating rejecting Absoluteness in favor of his theory of attitudes as having centered-worlds contents, that is, contents that vary in truth value not only with worlds but also with individuals. His argument was that since each god knows all absolute propositions, in order to explain why they still do not know which of the two they are, we need to appeal to non-absolute propositions.

Again Cappelen and Dever argue that this case is on a par with Frege cases:¹¹

what Lewis is trying to explain is how Zeus can know that *Zeus* is the god on the tallest mountain, but not know that *he* is the god on the tallest mountain. But this, of course, is just a special case of traditional Frege puzzles. (Cappelen and Dever, 2013, 99)

As for Messy Shopper, we might give the following Anti-Exceptionalist analysis of

Two Gods:

- (4) a. z believes $\langle z$ is on the tallest mountain, \ast Zeus is on the tallest mountain \ast \rangle
 b. z does not believe $\langle z$ is on the tallest mountain, \ast I'm on the tallest mountain \ast \rangle
 c. For any w , $\langle z$ is on the tallest mountain, \ast Zeus is on the tallest mountain \ast \rangle is true at w iff $\langle z$ is on the tallest mountain, \ast I'm on the tallest mountain \ast \rangle is true at w .

So this Anti-Exceptionalist claims that one can meet Frege's Constraint with respect to Perry cases by postulating a MOP, $\ast I \ast$, to differentiate thoughts like "Zeus lives on the tallest mountain" and "I live on the tallest mountain" when the thinker of both thoughts is Zeus.

One may of course be dissatisfied with this approach to Perry cases if one independently has misgivings about MOPs or absolute propositions or the idea that contents of attitudes are pairs of such things. Yet such misgivings are irrelevant to the topic under discussion here. The issue here concerns the claim that Perry cases demand nothing more of theories of attitudes than Frege cases. To complain that MOPs are theoretically unsatisfactory is not a way of vindicating the opposing Exceptionalist position. Correspondingly, the exercise of asking whether the theoretical demands of Perry cases go beyond those of Frege cases is a way of asking what exactly Perry cases show.

Moreover, I am not concerned to argue here that *any* way of handling Frege cases can also handle Perry cases. There may be approaches to the former that are well-motivated and maybe even preferable, given certain other theoretical commitments, for which Perry cases do present a distinct

11 Ninan (2016: 206) agrees.

problem. Rather, if what I argue here is on the right track, then at least some ways of dealing with Frege cases also apply to Perry cases. Still, even if Perry cases do not oblige one to abandon the Doctrine of Propositions, one might want to do so for other reasons.

To evaluate the Anti-Exceptionalist suggestion that since Perry cases can be handled in terms of first-person MOPs, they do not mandate a departure from the Doctrine of Propositions, we need to ask what is required of first-person MOPs to explain the relevant *de se* cases. I turn to this question in the next section.

3. Self-Reference and Non-Descriptiveness

In this section I point out two features that first-person MOPs arguably must have in order to explain Perry cases along with some other characteristics of first-person thoughts.

3.1. Self-Reference

The first feature of first-person MOPs to note is that such MOPs guarantee reference to oneself. Let us state this as follows:

Self-Reference

For all x and y , if x thinks of y under $*I*$, then $x = y$.

The reason $*I*$ must obey Self-Reference is that if it does not, the truth conditions of first-person thoughts will come out wrong.

Take Perry's (1993 [1977]) example of Heimson who has gone mad and thinks he is David Hume, also discussed by Lewis (1979).¹² He thinks to himself, "I wrote the *Treatise*." Heimson is wrong. What he thought is false. Heimson did not write the *Treatise*, Hume did. This means that, for the Anti-Exceptionalist, the propositional component of Heimson's thought involves Heimson, not Hume.

In other words, the Anti-Exceptionalist should analyze Heimson's thought as in (6) (where e is Heimson).

(6) $\langle e \text{ wrote the } \textit{Treatise}, *I \text{ wrote the } \textit{Treatise}* \rangle$

Moreover, many will agree that, even though there may be a sense in which Heimson thought about Hume (I return to this in 4.2), Heimson cannot think

(7) (where u is Hume).

(7) $\langle u \text{ wrote the } \textit{Treatise}, *I \text{ wrote the } \textit{Treatise}* \rangle$

Heimson cannot have a thought that is true if and only if Hume wrote the *Treatise* by thinking "I wrote the *Treatise*." Self-Reference secures this result.

¹² See also Ninan (2016, 88–89). For a related problem, see Barwise and Perry (1983, 148), and Stokke (in press-a) for relevant discussion.

Correspondingly, consider the following story:

Messy Heimson

Heimson is pushing his cart around the supermarket. Hume is in the same store, also shopping. Heimson notices that someone is making a mess. A clerk is heard announcing over the store's PA system, "Hume is not the one making a mess." The clerk is right. Hume is not making a mess. Heimson thinks, "Oh, good, I'm not the one making a mess." But for good measure he checks his cart. Everything is in order. Heimson feels relieved.

Clearly, in this case, what Heimson thought was true. Yet it was not true because Hume was not making a mess, but because Heimson himself was not making a mess. If the clerk had been wrong, and it really was Hume who was making a mess, Heimson's belief would still have been true.

Whatever one wants to say about *I*, the MOP involved in thinking things like "I'm making a mess" or "I wrote the *Treatise*," it must obey Self-Reference.

3.2. *Non-Descriptiveness*

Further, one way of stating the insight demonstrated by Perry cases is that, as Recanati (2012) puts it,

for any indexical α and non-indexical description 'the F ', it is always possible for the subject to doubt, or to wonder, whether α is the F [...].
(Recanati, 2012, 32)

I take it that by "it is always possible" here Recanati has in mind what we might formulate as "there are cases in which it is rational." That is, Perry cases show that, for instance, it can be rational to think "Now is not the time of the meeting," "This is not the road to Rome," "I'm not the person lost in the Stanford Library," and so on.

The point is not just the obvious one that it can be rational to think such things in situations where they are true and one has good reasons to believe them. Rather, Perry cases demonstrate that it can be rational to think such things even if they are false. That is, even if, for instance, "Now is the time of the meeting" is true or "I'm the person lost in the Stanford Library" is true, and so on.

Accordingly, for the first-person, Perry cases can be seen to illustrate that *I* is non-descriptive in the following broad sense:

Non-Descriptiveness

For any x and non-indexical F , x can rationally think $\langle x$ is not the F ,
*I'm not the F * \rangle

Moreover, it is a consequence of Self-Reference that when one thinks such things one is thinking about oneself. In other words, Non-Descriptiveness

implies that one can think about oneself as "I" even if one does not think that one is the *F*, for any non-indexical *F*.¹³

As suggested, this should be restricted to non-indexical descriptions. It is not clear that there are situations in which one can rationally think "I'm not the thinker of this thought," "I'm not here now," or the like, as long as the indexicals are read in the relevant way. However, it can be argued that, even so, nothing is gained by equating *I* with an indexical description.

For instance, consider the proposal that *I* should be understood as *the thinker of this thought*. That is, to have a thought τ of oneself as *I* is to think of oneself as the thinker of τ . Recanati rejects this interpretation of MOPs like *I*:

this move cannot support a descriptivist approach to indexical modes of presentation. What is needed to support such an approach is an objective, *non-indexical* description that provides the sense of the indexical. (Recanati, 2012, 33)

Yet he does not provide an argument for this rejection. I assume the motivation is the following. To suggest that when *x* thinks, for instance, "I'm French," she is thinking $\langle x$ is French, *The thinker of this thought is French* \rangle assumes that the demonstrative figuring in the MOP refers to $\langle x$ is French, *The thinker of this thought is French* \rangle itself. Clearly, though, thinking "the thinker of this thought is *F*" does not itself guarantee that "this thought" refers to one's own thought. You might think to yourself, "Someone thinks that Sue is making a mess. But the thinker of this thought is actually the one making a mess."

So, on this view, there must be a special demonstrative, call it "this+" that figures in first-person MOPs. For instance, one might postulate that *The thinker of this+ thought* always picks out the thinker herself. Indeed, this new MOP must obey Self-Reference in order to get the truth conditions of thoughts like "I wrote the Treatise" right. Yet, so understood, *The thinker of this+ thought* is not substantially different from *I*. There is no significant theoretical difference between the two proposals. For this reason, I will ignore indexical descriptions in what follows.

Given Self-Reference, Non-Descriptiveness means that *I* is a way that *x* thinks of herself that is not facilitated by any descriptive information she might associate with herself. A common way of putting this is to say that everyone is *acquainted* with themselves in a non-descriptive way.¹⁴ Acquaintance, in this sense, is the kind of relation that is standardly invoked to account for the way in which, for instance, perception can facilitate non-descriptive ways of thinking about things. As Recanati writes,

13 Since we are assuming that *x* is rational, it is safe to assume that if *x* thinks she is not the *F*, she does not think she is the *F*.

14 On this, see especially Recanati (1993, 72–73), (2012, 34–38).

in some cases, we are simply unable to properly describe the object that is given to us in experience: we don't know what it is, yet that does not prevent us from referring to it directly (without conceptual mediation) and e.g. wondering what it can be [...]. (Recanati, 2012, 29)

In such cases the subject thinks of an object under a non-descriptive MOP that is supported by an acquaintance relation.

Thoughts supported by acquaintance are typically known as *de re* (or *singular*) thoughts.¹⁵ Thinking *de re* about something does not exclude that one associates descriptive information with the relevant object. The kind of example Recanati describes in the quoted passage above is a limit case. For instance, if Ralph has seen a man in a brown hat sneaking around, he can think *de re* of that man in virtue of acquaintance. But still, he might associate descriptive information with him such as "wears a brown hat." Yet in this case, such information does not determine or constrain reference. If the hat is actually purple, Ralph's thought is still about that man. Rather, Ralph's thought refers to the man *de re* because that is the man he has been acquainted with through perception.

Correspondingly, it is natural to take *I* to be supported by an acquaintance relation. (I return to this in 4.3.) Yet we can note that all that is required is that the reference of *I* be determined relationally rather than satisfactorily, regardless of whether one thinks that this is underwritten by acquaintance or not. In particular, when *x* thinks of herself under *I*, her thought refers to *x* independently of any information she might associate with herself. For concreteness, I continue to talk of this kind of relationally in terms of acquaintance in what follows.

4. Three Exceptionalist Replies

In the last section I argued that to explain Perry cases, first-person MOPs need to be self-referential and non-descriptive, or relational. In this section I consider four potential Exceptionalist replies to understanding *de se* thoughts in terms of such MOPs.

4.1. *De Se vs. De Re*

Self-Reference and Non-Descriptiveness entail that when the messy shopper, *j*, comes to think "I'm making a mess," *j* is thinking about *j* in a non-descriptive, relational way. The first potential objection I want to comment on argues that this merely captures the *de re* sense in which *j* is thinking about *j* but does not explain the sense in which *j* is thinking *de se* about *himself*.

Perry (1993 [1979]) explicitly argued that *de se* beliefs cannot be analyzed as *de re* beliefs. However, Perry did not discuss the kind of view we

15 See Jeshion (2010) for a useful overview.

have outlined above according to which when *j* believes "I'm making a mess," he believes a proposition that is true if and only if *j* is making a mess under the MOP *I'm making a mess*. As I explain below, Perry's misgivings about understanding the *de se* as *de re* do not apply to this view.

In order to argue that the *de se* cannot be reduced to the *de re*, Perry focuses on a particular way of understanding *de re* beliefs, which Perry describes as a view on which the *de re* is understood in terms of *de dicto* beliefs. We can state this view as follows:¹⁶

De Re as De Dicto

x believes *de re* of *y* that *y* is *F* iff there is a concept α such that α fits *y* and *x* believes that α is *F*.

Here is Perry's objection to this view qua proposal for understanding the *de se*:

if this is our analysis of *de re* belief, the problem of the essential indexical is still with us. For we are faced with the same problem we had before. I can believe that I am making a mess, even if there is no concept α such that I alone fit α and I believe that α is making a mess. (Perry, 1993 [1979], 40–41)

I take this to be the same observation that we noted earlier when motivating Non-Descriptiveness. That is, I take it that by a "concept" here, we are to understand some descriptive information associated with the relevant object.

The most natural way of reading Perry's comment above is as the claim that I can believe "I'm making a mess" while disbelieving things like "The shopper with the torn sack is making a mess," even if I am the shopper with the torn sack. As such, *I*, as we have characterized this MOP, is not open to this objection. Indeed, when *j* thinks of *j* under *I* he is thinking directly of *j* without this being mediated by descriptive information, analogously to the standard way of understanding the relation between an ordinary proper name and its referent.

Having dismissed this construal of *de re* beliefs, Perry (1993 [1979], 42) considers a more intuitive view on which "I believed of John Perry that he was making a mess." Unsurprisingly, Perry dismisses this proposal, too:

Saying that I believed of John Perry that he was making a mess leaves out the crucial change, that I came to think of the messy shopper not merely as the shopper with the torn sack, or the man in the mirror, but as *me*. (Perry, 1993 [1979], 42)

Yet this is not the view we have outlined above, either. While it is true to say that, on that view, *j* believes of *j* that he is making a mess, he does so while thinking of *j* as "I." The latter feature of *j*'s state of mind is what we analyze by saying that *j* thinks of *j* under the MOP *I*.

¹⁶ Cf. Perry (1993 [1979], 40). With some caveats, this is a version of the view of *de re* beliefs pioneered by Kaplan (1969).

4.2. Transparency

The second Exceptionalist response I want to consider argues that the Anti-Exceptionalist who understands *de se* thoughts in terms of first-person MOPs has problems accounting for a particular kind of *transparency* that is characteristic of *de se* thoughts.

There is an intuitive sense in which one can think things like "Hume is making a mess" or "The shopper with the torn sack is making a mess" without being able to identify the referent. For instance, thinking, or even knowing, "The shopper with the torn sack is making a mess" does not guarantee that one is able to pick out the messy shopper or demonstrate them if one was presented with them. As we might say, one can think "The shopper with the torn sack is making a mess" without knowing who the shopper with the torn sack is. The same holds for "Hume is making a mess."

This intuition might be made concrete by claiming that MOPs like *The shopper with the torn sack* and *Hume* are non-transparent in the following sense:

Non-Transparency

A MOP $*m*$ is non-transparent iff for any x and y , x can think $\langle y$ is F , $*m$ is F^* \rangle without being in a position to identify y .

On the other hand, *I* is arguably transparent. If you think "I'm making a mess," there is no sense in which you might not be in a position to identify yourself, or know who you are.

This idea is reminiscent of one way of understanding the phenomenon known as *immunity to error through misidentification*. For instance, Coliva (2003) writes,

Consider, for instance, the case of a person who uses 'me' to refer to someone, different from herself, she sees reflected in the mirror. If there is a sense in which she can take that person to be herself, then there is a sense in which her use of 'I' can refer to that person. I think there is such a sense. (Coliva, 2003, 426)

Coliva concludes that "I" obeys what she calls "The Real Guarantee:"

The Real Guarantee (at the level of language): the comprehending use of 'I' guarantees that the speaker knows which person is its *semantic* referent. (Coliva, 2003, 428)

Take a version of the case Coliva mentions. Lisa sees someone reflected in a shop window. She thinks it is herself. Noticing that the person has a white stain on her shirt, Lisa thinks, "I have a white stain on my shirt." However, it was not herself but Anna she saw reflected. One might argue that there is a sense in which Lisa intended to refer to Anna in this case, since Anna is the person she saw reflected. Yet clearly what Lisa said is false, since Lisa does not have a white stain on her shirt.

So one can take Anna to be the speaker referent of Lisa's use of "I," while Lisa is the semantic referent. This accords with Self-Reference, as we have understood this notion above. We said that when Heimson thinks "I wrote the *Treatise*" he cannot be thinking that Hume wrote the *Treatise*, even though he thinks he is Hume. If one likes, one can insist that Heimson speaker-refers to Hume with this thought, but there is no sense in which what Heimson thought was true. That is, the truth conditions of his thought undeniably involve Heimson (the semantic referent), not Hume.

In other words, the suggestion is that when x thinks of x under $*I^*$, she is in a position to identify herself as the (semantic) referent of her thought. $*I^*$ is transparent, as opposed to MOPs like $*That\ guy^*$, $*Hume^*$, $*The\ shopper\ with\ the\ torn\ sack^*$ and many others. The challenge concerns how the Anti-Exceptionalist can account for this feature of *de se* thoughts.

A first pass is to acknowledge that $*I^*$ is transparent in the following way:

Transparency

For any x , if x thinks $\langle x$ is F , $*I$ am F^* \rangle , x is in a position to identify x .

But what can the Anti-Exceptionalist mean by " x is in a position to identify x "? First, consider Non-Transparency. The observation is that you can think, for instance, "The shopper with the torn sack is making a mess" without being in a position to identify the shopper. To identify the shopper, intuitively, is to think something like, "The shopper with the torn sack is so-and-so," that is, to have a thought like $\langle j = j$, $*The\ shopper\ with\ the\ torn\ sack\ is\ John\ Perry^*$ \rangle , or the like. In other words, a natural way of understanding what it means to identify someone is be able to think of them as one side of an informative identity. Given the apparatus of MOPs, this means to be able to think of them under two different MOPs, as in $*m_1$ is m_2^* .

If this is right, then Transparency must be understood as the claim that when one thinks "I am F ," one is able to think of oneself as one side of an informative identity of the form "I am so-and-so." Hence, Transparency can be re-stated as

Self-Identification

For any x , if x thinks $\langle x$ is F , $*I$ am F^* \rangle , there is an MOP $*m^*$ such that x is in a position to think $\langle x = x$, $*I$ am m^* \rangle

The question is what m in Self-Identification could be.

We already know that m cannot be an MOP like $*The\ shopper\ with\ the\ torn\ sack^*$, $*John\ Perry^*$, or $*That\ guy^*$. For instance, you can think "I'm hungry" without thinking "I'm the shopper with the torn sack," even if you are the shopper with the torn sack, or you can think "I'm about to be hit by snow" without thinking "I'm that guy," even if you are unwittingly demonstrating yourself.

Indeed, it seems that the only candidate for m in Self-Identification is $*I*$ itself. So, according to this argument, the only way the Anti-Exceptionalist can account for Transparency is by noting that whenever x thinks "I am F " she is able to think $\langle x = x, *I \text{ am } I^* \rangle$. Yet, so this objection goes, this is trivial, and does not explain the sense in which "I" is special in that when one thinks "I am F ," one knows who one is thinking about, or alternatively, one is in a position to identify oneself as the (semantic) referent, in contrast with thinking things like "Hume is F ." Indeed, the identity $*I \text{ am } I^*$ is not informative, since the same MOP appears on each side.

The Anti-Exceptionalist can respond to this challenge by demanding a more substantial sense of what is meant by the claim that one is always in a position to identity oneself as "I." Indeed, she can insist that there is only a trivial sense in which this is right. Namely, when you think "I am F " the referent of your thought is your-self, as per Self-Reference. In other words, you are guaranteed to know who the referent of "I" is because you are identical with that person. Yet the Exceptionalist should exhibit some phenomenon or data point about *de se* thinking that is not accounted for by understanding the *de se* in terms of first-person MOPs.

4.3. Pyrrhic Anti-Exceptionalism?

Another way of reacting to the view we have outlined here is to accept that it is a way of preserving Two-Place and Absoluteness by postulating first-person MOPs as an implementation of Frege's Constraint in the face of the *de se* phenomena. Yet even if one agrees that it has been shown that Perry cases are on a par with Frege cases in that they do not demand more than the Doctrine of Propositions allows, one might argue that what Perry cases demand of $*I*$ still corroborates the claim that *de se* thoughts form a substantially distinct category of thoughts.

If this is right, the Anti-Exceptionalist may be said to have won a merely pyrrhic victory. Below, I consider Self-Reference and Descriptiveness in turn. As I explain, none of these characteristics of first-person MOPs are exceptional enough to render the Anti-Exceptionalist an Exceptionalist in disguise.

First, consider Self-Reference. This feature of $*I*$ means that first-person MOPs are guaranteed to refer, and moreover are guaranteed to refer to the subject of the relevant thought. The former feature is not different from many other MOPs. For instance, you might believe both "8 is smaller than 9" and "the square root of 64 is larger than 9" if you do not realize that 8 is the square root of 64. Yet $*\text{the square root of } 64^*$ is guaranteed to refer. The fact that $*I^*$ is guaranteed to refer to x is arguably not an exceptional feature either. After all x is an entity to which one can refer, and in particular, an entity to which x can refer.

Second, as we saw (in 3.2), Non-Descriptiveness is likewise a feature of many other MOPs. Demonstrative, perception-based MOPs like *that man over there* and so on, are ways of thinking directly about things. In particular, as we noted, non-demonstrative MOPs are characterized by their reference being determined relationally, rather than satisfactionally.

Acquaintance is a general way of understanding relational MOPs of this kind. If one can think non-descriptively about other people, it is not surprising that one can think non-descriptively about oneself. Moreover, as we said, any acquaintance relation that you can bear to something not identical to yourself is open to Perry cases. In other words, *I* must be supported by a relation that you can only bear to yourself.

What relation could this be? One suggestion is: identity. Indeed, Perry (2002) has suggested that thinking about oneself is facilitated by a mental file supported by the relation of self-identity.¹⁷ So, the claim would be that if one accepts that the reference of some MOPs is determined relationally, for instance, when thinking about something perceived visually, there is nothing exceptional in the suggestion that identity is a relation that can facilitate thinking in this way about someone, namely oneself.

In particular, one can follow Lewis (1999 [1983]), Recanati (2012), and others in broadening the notion of acquaintance to relations that are, to use Recanati's term, "epistemically rewarding." That is, one can accept that there are many kinds of relations that can facilitate the kind of direct, non-descriptive way of thinking about something that is captured by relational MOPs. Lewis writes,

There are the relations that someone bears to me when I get a letter from him, or I watch the swerving of a car he is driving, or I read his biography, or I hear him mentioned by name, or I investigate the clues he has left at the scene of his crime. In each case, there are causal chains from him to me of a sort which would permit a flow of information. Perhaps I do get accurate information; perhaps I get misinformation, but still the channel is there. I shall call such relations as these *relations of acquaintance*. (Lewis, 1999 [1983], 380–381)

As described by Lewis here, the relevant relations involve "causal chains." Yet one might wonder to what extent self-identity could be a relation such that there may be causal chains from one relatum to the other, that is, from oneself to oneself.

Recanati writes,

In virtue of being a certain individual, I am in a position to gain information concerning that individual in all sorts of ways in which I can gain information about no one else, e.g. through proprioception and kinaesthesia. (Recanati, 2012, 36)

¹⁷ See also Recanati (2012, 36–37).

One way of understanding this picture is that self-identity enables a privat acquaintance relation of the Lewisian kind, which allows for causal flow of information, in the sense that proprioception (or kinaesthesia) is facilitated causally by neurons interacting with the central nervous system.

Along these lines, one might argue that Self-Reference and Non-Descriptiveness are simply features of this kind of acquaintance relation based on self-identity, and moreover, not features that make *I* substantially different from a host of other MOPs based on acquaintance relations.

References

- Babb, M. (2016). The essential indexical of intentional action. *The Philosophical Quarterly*, 66(264), 439–457.
- Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge, MA: MIT Press.
- Bermudez, J. (2017). Yes, indexicals really are essential. *Analysis*, 77, 690–694.
- Burge, T. (2004). Memory and persons. *Philosophical Review*, 113, 289–337.
- Cappelen, H., & Dever, J. (2013). *The inessential indexical – on the philosophical insignificance of perspective and first person*. Oxford and New York: Oxford University Press.
- Castañeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, 8, 130–157.
- Castañeda, H.-N. (1975). *Thinking and doing*. Dordrecht: Reidel.
- Coliva, A. (2003). The first person: Error through misidentification, the split between speaker's and semantic reference, and the real guarantee. *Journal of Philosophy*, 100(8), 416–431.
- Devitt, M. (2013). The myth of the problematic *De Se*. In A. Capone & N. Feit (Eds.), *Attitudes de se: Linguistics, epistemology, metaphysics*. Stanford: CSLI Publications.
- Douven, I. (2013). The epistemology of *De Se* beliefs. In A. Capone & N. Feit (Eds.), *Attitudes de se: Linguistics, epistemology, metaphysics*. Stanford: CSLI Publications.
- Evans, G. (1982). *The varieties of reference* (J. McDowell, Ed.). Oxford: Oxford University Press.
- Frege, G. (1956 [1918]). The thought: A logical inquiry (A.M. & M. Quinton, Trans.). *Mind*, 65(259), 289–311.
- Jeshion, R. (2010). Introduction. In R. Jeshion (Ed.), *New essays on singular thought* (pp. 1–40). Oxford and New York: Oxford University Press.

- Kapitan, T. (1995). Intentions and self-referential content. *Philosophical Papers*, 24, 151–166.
- Kaplan, D. (1969). Quantifying in. In D. Davidson & J. Hintikka (Eds.), *Words and objections: Essays on the work of W. V. O. Quine* (pp. 206–242). Dordrecht, Holland: Reidel.
- Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford and New York: Oxford University Press.
- Lewis, D. (1979). Attitudes *De Dicto* and *De Se*. *Philosophical Review*, 88, 513–43.
- Lewis, D. (1999 [1983]). Individuation by acquaintance and by stipulation. In *Papers in metaphysics and Epistemology* (pp. 373–402). Cambridge and New York: Cambridge University Press.
- Magidor, O. (2015). The myth of the *De Se*. *Philosophical Perspectives*, 29, 249–283.
- McGinn, C. (1983). *The subjective view: Secondary qualities and indexical thoughts*. Oxford: Clarendon Press.
- Millikan, R. (1990). The myth of the essential indexical. *Noûs*, 24(5), 723–734.
- Morgan, D. (2019). Temporal indexicals are essential. *Analysis*, 79(3), 452–461.
- Ninan, D. (2010). *De se* attitudes: Ascription and communication. *Philosophy Compass*, 5(7), 551–567.
- Ninan, D. (2013). Self-location and other-location. *Philosophy and Phenomenological Research*, 87(2), 301–331.
- Ninan, D. (2016). What is the problem of *De Se* attitudes? In M. García-Carpintero & S. Torre (Eds.), *About oneself: De se thought and communication* (pp. 86–120). Oxford and New York: Oxford University Press.
- Owens, D. (2011). Deliberation and the first person. In A. Hatzimoysis (Ed.), *Self-knowledge* (pp. 261–278). Oxford and New York: Oxford University Press.
- Perry, J. (1993 [1977]). Frege on demonstratives. In *The problem of the essential indexical and other essays* (pp. 3–32). Oxford and New York: Oxford University Press.
- Perry, J. (1993 [1979]). The problem of the essential indexical. In *The problem of the essential indexical and other essays* (pp. 3–52). Oxford and New York: Oxford University Press.
- Perry, J. (2002). *Identity, personal identity, and the self*. Indianapolis: Hackett.
- Prosser, S. (2015). Why are indexicals essential? *Proceedings of the Aristotelian Society*, 115(3), 211–233.

Recanati, F. (1993). *Direct reference – from language to thought*. Oxford: Blackwell.

Recanati, F. (2012). *Mental files*. Oxford and New York: Oxford University Press.

Salmon, N. (1986). *Frege's puzzle*. Cambridge, MA: MIT Press.

Stalnaker, R. (1999 [1981]). Indexical belief. In *Context and content* (pp. 130–149). Oxford and New York: Oxford University Press.

Stokke, A. (in press-a). Features of referential pronouns and indexical presuppositions. *Inquiry*.

Stokke, A. (in press-b). Navigation and indexical thought. *Erkenntnis*.

Velleman, D. (1996). Self to self. *Philosophical Review*, 105, 39–76.

Weber, C. (2014). Indexical beliefs and communication: Against Stalnaker on self-location. *Philosophy and Phenomenological Research*, 90, 640–663.



FOR THEIR EYES ONLY

Abstract: *When and why do we need the indexical 'I'? Perry (1979) thinks that 'I' is an essential ingredient to the explanation and prediction of action. We need 'I' to classify the kind of belief that causes an agent to produce a new action. In his view, classifying the agent's belief in terms of 'I' makes sense because, when asked to explain her behavior, the agent will be disposed to say 'I'. Here, we argue that this dispositional assumption is problematic. The disposition to act according to an essential indexical and the disposition to make a speech act that contains 'I' are two very different things that are not always related. The most common circumstances in which we need to say 'I' to explain ourselves involve other agents: when we coordinate in joint actions or when we engage in self-talk to save face before the eyes of potential witnesses. Finally, we revisit Perry's famous messy shopper anecdote to offer a better story about why its main character is disposed to say 'I'.*

Keywords: *Essential indexical; Speech acts; Dispositions; Coordination; Self-talk.*

1. Introduction

Remember the messy shopper? He is the main character of the anecdote told by Perry (1979) about a time he followed a trail of sugar on a supermarket floor, pushing his cart around in pursuit of the shopper who was making a mess, only to find out that the sugar was coming from the torn sack inside his own cart. As Perry tells us, it was after he realized that *he*, John Perry, was the messy shopper that he stopped to fix the torn sack. According to his account, then, having the belief 'I am making a mess' caused Perry to change his conduct. He concluded that the 'I' in the acquired belief is an essential ingredient to the explanation of his new action, since, in replacing it by alternative designations – for instance, the definite description, 'the only bearded philosopher in a Safeway store west of the Mississippi', or the proper name, 'John Perry' –, we seem to no longer be able to account for his change in behavior – from pursuing the messy shopper to fixing the torn sack in his own cart.

The thesis of an essential indexical ingredient to the explanation of action also appears in a second story from the same text. It is the example of the tardy professor that abandons his office hurriedly at noon, after realizing that the department meeting scheduled for that day started at noon. Here too,

according to Perry, we need an indexical belief like ‘The department meeting starts at noon and noon is *now*’ to explain the change in the professor’s behavior.

In both cases, Perry describes the characters as “ready to say” or “disposed to utter” indexical sentences with the words ‘I’ and ‘now’.¹ This claim, call it *the dispositional assumption*, implies a relation between indexical beliefs and utterances with indexical expressions that, as we will argue in what follows, creates some difficulties for the position. Our point, essentially, is that the disposition to utter an indexical sentence is a disposition to perform a speech act and that changes in conduct need not be accompanied by speech acts that explain them. In other words, to make sense of the messy shopper’s action of fixing the torn sack as involving the first-person perspective (the essential ‘I’), we do not need to assume that he is disposed to explain or predict his own behavior by uttering the word ‘I’. Or at least Perry does not offer an explanation as to *why* we should assume that. In the best-case scenario, this appeal to a disposition to speak needs an additional story. Here, we offer such an additional story by considering two kinds of situations in which agents are disposed to utter ‘I’ to explain their own behavior: coordination in joint activities and cases of (save face) self-talk (Goffman 1978). We argue, in particular, that the best way to explain the messy shopper’s assumed disposition to say ‘I am making a mess’ is to take it as case of save face self-talk.

We begin our discussion, in the next section, by pointing out that the original formulation of the thesis of the essential indexical involves a dilemma. Though it purports to establish the ineliminability of indexicals in the explanation of action, the position ends up committing to a classificatory view that entails a third-person perspective in which ‘I’, ‘here’ and ‘now’ actually have no use. In section 3, we discuss two possible ways out of this dilemma: by appealing to the argument of the language of thought or to the disposition to utter indexical expressions. Perry seems to favor the last one. In sections 4 and 5, we hold that a dispositional solution to the dilemma requires an additional story about why agents are disposed to utter ‘I’, ‘here’ or ‘now’. We first consider situations of joint activities and then of self-talk, before we revisit the anecdote of the messy shopper.

2. A dilemma from Kaplan

In his *Demonstratives* (1989), David Kaplan has two sorts of motivation for a special semantics for indexicals. The first one is based on the truth-conditional profile of sentences with such expressions: how they behave in

1 “Clearly and correctly, we want the tardy professor, when he finally sees he must be off to the meeting, to be ready to say „I believe that the time at which it is true that the meeting starts now is now.”” Perry 2000: 45. “We expect all good-hearted people in the state that leads them to say „I am making a mess“ to examine their grocery carts, no matter what belief they have in virtue of being in that state” Perry 2000: 39–40.

modal, past, future and counterfactual sentences² and what inferential patterns they authorize.³ This is accounted for by his Logic of Demonstratives. The second kind of motivation concerns the cognitive significance of mental states represented by such sentences – following Kaplan himself, we may call it the “epistemological motivation”⁴. Actually, the apparatus developed for the semantic motivation is used to solve some puzzles related to the epistemological motivation, concerning the nature of our thoughts: in broad terms, the cognitive significance of a thought, in Kaplan’s account, corresponds to its character.⁵

After suggesting a solution to Frege’s Puzzle within his framework, Kaplan tries a “different line of argument”:

Now instead of arguing that character is what we would ordinarily call cognitive significance, let me just ask why we should be interested in the character under which we hold our various attitudes. Why should we be interested in that special kind of significance that is sensitive to the use of indexicals; ‘I’, ‘here’, ‘now’, ‘that’ and the like? John Perry, in his stimulating paper “Frege on Demonstratives”, asks and answers this question.⁶

And then he quotes Perry:

Why should we care how someone apprehends a thought, so long as he does? I can only sketch the barest suggestion of an answer here. We use senses to individuate psychological states, in explaining and predicting action. It is the sense entertained and not the thought apprehended that is tied to human action.⁷

The point of view of the agent, including her grasp of what she thinks about here and now, is presented by Perry then as ineliminable in the explanation of action. However, there is an important difference between their aims: while Kaplan is talking about indexical *expressions*, Perry is interested in the account of our *beliefs* and *actions*. As he puts it in his 2020 book, he

<...> generalized Kaplan’s theory to beliefs. Just as there were sentences that could be used to say different things, by different people or at different times, there were *ways of believing* that constituted believing different things for different people or at different times.⁸

2 Kaplan 1989: 498–500, 512–514.

3 Kaplan 1989: 522.

4 Kaplan 1989: 529.

5 Kaplan 1989: 529–540.

6 Kaplan 1989: 532.

7 Perry 2000: 23, also quoted in Kaplan 1989: 532.

8 Perry 2020: 10.

We are then after the *roles* of indexicals, to use Perry's vocabulary, not the meanings of indexical words: "What correspond to indexicals in our beliefs are not words with a special conventional meaning, but notions and ideas that are tied to what I call "epistemic and pragmatic roles."⁹

Roles "provide a way of organizing information that comes naturally to humans and is reflected in many ways in language."¹⁰ It is a way to gather information about an object that is the current focus of attention which corresponds to an indexical belief. This object may be called "the index."¹¹ In order to track it in a scene, or to think about one's current location and time, or, indeed, to have a first-person perspective on a given situation, there is no need of a linguistic expression.¹² But mental indexicals, identified by their role, have a clear similarity with the meanings of indexical words. The disposition to utter an indexical sentence is a way to explain this similarity. This sort of account, however, is less than smooth.

Let us take a look at the messy shopper again: he believed the same singular proposition before and after realizing that he was making a mess – there is no change in *what* he believed. Saying that he believed the same thing all along is a way to classify his mental states that is useful for some purposes, for instance, if we want to keep track of the truth-conditions of his beliefs. Singular propositions are abstract structures useful to classify different agents, or the same agent at different moments of time, according to what is believed. Indexicals have a part to play here: in order to have the same belief across relevantly different contexts, one must believe it in different ways. Indexicals track precisely the different ways in which one may believe the same singular proposition in different contexts. This is Frege's point about demonstratives.¹³ However, in order to explain behavior, these different ways to believe classify agents differently, as the tardy professor's anecdote shows:

As time passes, I go from the state corresponding to "The meeting will begin" to the one corresponding to "The meeting is beginning" and finally to "The meeting has begun." All along I believe of noon that it is when the meeting begins. But I believe it in different ways. And to these different ways of believing the same thing, different actions are appropriate: preparation, movement, apology.¹⁴

9 Perry 2007: 518.

10 Korta and Perry 2011: 28.

11 "Sometimes when we are dealing with a lot of facts about numerous inter-related objects, one object will take center stage for a period of time, during which we focus on which objects stand in various relations to it, or, as we say, play various roles relative to it. We call this object 'the index.'" Korta and Perry 2011: 29.

12 For a distinction between indexicals in language and in thought, see Recanati 2012: 57–67

13 Frege 1956: 296 *apud* Perry 2000: 8.

14 Perry 2000: 49.

Sentences with indexicals classify ways of believing, whether we are interested in classifying the agent as having the same thought across contexts, or as behaving in different ways while having in mind the same thought: “We use sentences with indexicals or relativized propositions to individuate belief states, for the purposes of classifying believers in ways useful for explanation and prediction”.¹⁵

At first, one might feel tempted to say that sentences with indexicals are not *what* is believed, but *how* one believes. However, sentences are not mental states. We need to classify different ways of believing because different ways of having the same thought may lead to different actions. Actually, indexical sentences are useful for classificatory purposes, that is, to explain and predict action, as Perry says, precisely by tracking these ways of believing. How?

It is not entirely clear how we may hold this classificatory view of indexicals and, at the same time, claim that the words ‘I’, ‘here’, and ‘now’ are essential to the explanation of action, since the classificatory view makes sense only from a third-person point of view. The agent herself does not classify her own states in order to explain her behaviour (or at least not for the sake of explanation, as we will see below), and certainly not to predict her own actions. When words are used to classify “believers in ways useful for explanation and prediction”, there does not seem to be a word ‘I’, a word ‘now’, nor a word ‘here’.

Well, this is a bit far-fetched. A speaker may use ‘now’ and ‘here’ to talk about a context other than the context in which she finds herself. Consider this narrative of Alexander the Great’s conquests, around 331 BC:

Alexander now occupied Babylon, city and province; Mazaeus, who surrendered it, was confirmed as satrap in conjunction with a Macedonian troop commander, and quite exceptionally was granted the right to coin. As in Egypt, the local priesthood was encouraged. Susa, the capital, also surrendered, releasing huge treasures amounting to 50,000 gold talents; here Alexander established Darius’s family in comfort. Crushing the mountain tribe of the Ouxians, he now pressed on over the Zagros range into Persia proper and, successfully turning the Pass of the Persian Gates, held by the satrap Ariobarzanes, he entered Persepolis and Pasargadae.¹⁶

In this paragraph, ‘now’ occurs twice, and ‘here’, once. They refer to Alexander’s time and location, and not to the writer’s. There is a shift in the context of interpretation, and ‘here’ and ‘now’ are used to explain Alexander’s actions.¹⁷ That only goes to show that words are used in different ways, and, in this use, they are irrelevant to Perry’s argument: the writer is certainly not engaged in

15 Perry 2000: 47–48.

16 Extracted from <https://www.britannica.com/biography/Alexander-the-Great>. Accessed in May, 17th, 2021.

17 For more on this kind of context shifting, see Recanati 2000: 211–258; Perry 2001: 76–80.

any action in 331 BC¹⁸. When they seem to be essential to the explanation of action, as in the stories of the messy shopper and the tardy professor, ‘I’, ‘here’ and ‘now’ seem to have no use. To sum up, the dilemma is that there is no room for ‘I’, ‘here’ and ‘now’ in a classificatory view of indexicals, and the argument is supposed to establish, precisely, the ineliminability of indexicals. There seems to be two alternative paths.

3. Placing indexicals

The first way out is to associate the linguistic character of an expression with a psychological mode of presentation. In section XX of his *Demonstratives*, Kaplan asks how we can make an indirect discourse report of an indexical belief, or a report of an utterance of a sentence containing an indexical expression. His answer is that when I say I., in reporting, for example, ‘My God! It’s my pants that are on fire’, I am reporting a belief that the agent has under the character of ‘I’.

1. He thinks that his pants are on fire,

More, in Kaplan’s words, “<...> when you wonder, ‘Is that me?’ it is correct to report you as having wondered whether you are yourself. These transformations are traced to the indexical form of your inner discourse rather than to any particular referential intentions.”¹⁹

So, sentence 2 is a report of an indexical sentence in your “inner direct discourse”, 3:

2. You wonder whether you are yourself.
3. Is that me?

Kaplan expects this move to take care of the relational sense of propositional attitude reports discussed by Quine (1956)²⁰, at least as far as indexicals are concerned. In his account, the relational sense of the report of an indexical belief is the indirect discourse counterpart of such a direct reference device in the “inner direct discourse” of the subject whose belief is reported. Such a move would avoid Quinean worries about such constructions²¹, since there

18 The point here is that utterances of indexical words, like ‘I’, ‘here’ and ‘now’, do appear in explaining actions (in the sense of rationally reconstructing, justifying or narrating them and etc.), but “outside” of the contexts in which such actions are produced.

19 Kaplan 1989: 554.

20 For Quine, there are two readings (which correspond to different reports of states of mind) of ambiguous natural language sentences containing psychological verbs, like ‘believe’: notional and relational. When the relational reading is true, there is someone to whom the belief is related, via the token relational state of mind in question (Crawford 2008). The same does not hold for the notional reading.

21 The worry here is that relational reports are not referentially transparent – as shown in the example (from the same text) of Ralph believing of Bernard J. Ortcutt that he is both a spy and the community’s pillar –, so the singular term is not purely referential (as

is no “near-contrary” in the vicinity of 3, given that a subject cannot think of herself using an indexical and fail to recognize that she is thinking of herself.

This solution could also solve our problem if we assumed that an indexical that classifies a mental state from a third-person point of view has a counterpart in the language of thought of the agent. As Kaplan puts it, “if this is correct, and if indexicals are featured in the language of thought, then we have a solution to Quine’s puzzle.”²² But he is not entirely convinced of the truth of the antecedent, for at least three reasons:

First, there is no real syntax of the language of thought. <...> Second, in containing images, sounds, odors, etc., thought is richer than the language of report. <...> Third, thought ranges from the completely explicit (inner speech) to the entirely implicit (unconscious beliefs which explain actions) and through a variety of occurrent and dispositional forms.²³

One may be more optimistic about the syntax of the language of thought, and that responding to the first worry will take care of the other two.²⁴ Well, maybe so, but this is no answer to the dilemma about the dispositional assumption.

Before taking up this problem, let us bear in mind that Perry himself is not entirely comfortable with reading the structure of thought off the structure of language. Indeed, for him, at least as far as indexicals are concerned, this is a misinterpretation that follows from two mistaken assumptions: that for a sentence to be used to classify belief states, it should be the object of the belief, and that since indexicals “*seemed* essential to *expressing* certain thoughts,” they are necessary “for *having* those thoughts.”²⁵ As we saw in the previous section, in order to explain our actions, we need to track the causal roles of beliefs, and not the meanings of words.

As we also saw in the previous section, in Perry, indexicals enter the story via the dispositional assumption:

That we individuate belief states in this way doubtless has something to do with the fact that one criterion for being in the states we postulate – at least for articulate, sincere adults – is being disposed to utter the indexical sentence in question.²⁶

attested via substitution *salva veritate*), and we cannot conclude any relation between the subject having the belief and the object of the belief – e.g., Ralph and Bernard J. Ortcutt.

22 Kaplan 1989: 553.

23 Kaplan 1989: 556; on the second reason, see Kaplan 1968: 201.

24 In this respect, see, in particular, Corazza 2004.

25 Perry 1997: 19*n*.

26 Perry 2000: 39–40.

The fact that an agent is disposed to utter a sentence with an indexical in certain circumstances explains why her action is classified by that sentence. Should the messy shopper decide to say anything as he realizes that he is making a mess himself, he will utter a sentence like 4:

4. *I am making a mess!*

But is it correct to say that the messy shopper is disposed to utter 4 as he rearranges his sack in his cart?

4. When are we disposed to utter an indexical sentence? A coordination view

When are we disposed to utter a first-person sentence *as we act*? Here is a situation in which the utterance of a first-person sentence seems natural:

Suppose that Heinrich and Andrea are going for a walk together. <...> Suppose now that Heinrich suddenly claps his hands to his brow and says “Oh No!” and, without further ado, starts walking rapidly away from Andrea. Andrea may not be disappointed that he has gone. Barring special background understandings, however, she will understand that <...> the manner of going involved a *mistake*. It is a mistake by the fact that they were walking together. Heinrich’s walking away will not have this effect if he first says to Andrea “I’ve just remembered I am supposed to be at the doctor’s office in ten minutes!” and receives the reply “Go quickly, then!”²⁷

In this story, Heinrich says 5 to Andrea:

5. I’ve just remembered I am supposed to be at the doctor’s office in ten minutes!

He is certainly not predicting his action. Is he explaining it to Andrea? Maybe so, but why? He certainly does not explain each of his actions to his friend. In Margaret Gilbert’s account, Heinrich is committed to continue in his joint activity with Andrea, and his utterance of 5 is a way to ask permission to leave this activity, not simply to inform her of his intention.²⁸

There are two different acts, requiring distinct accounts: the act of walking away from Andrea and the act of saying something to her. Gilbert provides an explanation for the speech act of uttering 5: he explains why he has to walk away from her because he has assumed the commitment to have a walk with her, and breaking such commitment requires an explanation. The use of a first-person pronoun is explained by its role in the coordination of a joint action.

27 Gilbert 2014: 24–25.

28 Gilbert 2014: 25.

Does this account extend to Perry's messy shopper story? Well, if he is alone, he does not have to coordinate his actions with anybody. Probably, if the messy shopper decided to say anything, he would utter something like 4. But why should he say anything? Is he disposed to utter a sentence describing each of his actions? Addressed to whom? And why? The disposition to utter a sentence is the disposition to perform a certain kind of speech act, and is quite independent of the agent having an indexical sentence in his inner discourse. After all, even if having a thought is having a sentence in the inner discourse, no one is disposed to utter every sentence that corresponds to her current object of thought.

To bring home this point, let us take a look at John Campbell's remarks on the first-person perspective. Here is his analysis of a story by John Austin. One day, Austin decides to kill his donkey: he goes to the field, sees his donkey and shoots it. Unfortunately, he shot his neighbour's donkey instead. Here is Campbell's description of what happened (let us keep the same agent throughout the story – Campbell is speaking on behalf of Austin-the-shooter):

When I make a mistake, take a careful aim and fire at your donkey, I will say 'I am shooting my donkey', and this judgment really does involve an error of identification.²⁹

Why did he utter this sentence? Addressed to whom? Does he describe to himself every action he is engaged in ('I'm reading a paper', 'Now, I will have a coffee' etc.)? Of course, there is a wrong judgment about the identity of the donkey that explains his action, but that doesn't mean we need to postulate that a first-person sentence was or would have been actually uttered. Contrast Campbell's story with Austin's:

You have a donkey, so have I, and they graze in the same field. The day comes when I conceive a dislike from mine. I go to shoot it, draw a bead on it, fire: the brute falls in its tracks. I inspect the victim, and find to my horror that it is *your* donkey. I appear on your doorstep with the remains and say – what? 'I say, old sport, I'm awfully sorry etc., I've shot your donkey *by accident*'?³⁰

He only says something as he appears at his neighbour's doorstep and apologizes for his awful mistake. Of course, apologizing for shooting the donkey and shooting the donkey are two different acts. The circumstances in which the dispositional solution holds are not that of an agent realizing that the moment in which she finds herself is the time for her to do something, but the circumstances in which, on top of that, she is disposed to say to an audience that it is time for her to do something, or that she has done something. One may be disposed to shoot a donkey silently, and have no disposition for apologizing for that.

29 Campbell 2003: 162.

30 Austin 1979: 185*n*.

The distinction is even more striking in a second example from Campbell:

But this extreme view [that the referent of ‘I’ is a construct synthesized around uses of the first person] does not seem to be correct, as we can see if we reflect on, for example, the multiplicity of bases on which we make first-person judgements. There are many different bases that we use. There is my knowledge of what I am doing or am about to do, and my memory of my past life. There are such elementary phenomena as my sense of balance exercised when I think ‘I am about to fall over.’³¹

However, having a certain thought in mind is not the same thing as having in mind a sentence that expresses that thought. Think of the “swinging room” experiment:

Lee and Aronson placed 13- to 16-month-old toddlers in a “swinging room”. In this room, the floor was stationary, but the walls and ceiling could swing toward and away from the toddler. [...] Notice that this pattern is similar to the optic flow that occurs when moving forward, as when you are driving through a tunnel. Because the flow is associated with moving forward, it creates the impression in the observer that he or she is swaying forward. This causes the toddler to sway back to compensate.³²

The toddler thinks she is about to fall over, but she has no disposition to utter ‘I am about to fall over’: children start to understand personal pronouns by the age of 18-months.³³ The experiment addresses the agent’s perspective on a scene, but the agent having a perspective is not equivalent to her disposition to utter a first-person sentence describing the scene from her vantage point. Of course, Campbell knows this and his analysis requires no sentence in the agent’s mind – except as perhaps a *façon de parler*.³⁴

To recap then, the dispositional solution as applied to the messy shopper example seems incomplete without an additional story about why our character would need to perform a speech act aside from simply fixing the sugar sack in his cart. Since the messy shopper is not in a joint activity with a fellow shopper and he has no audience, why would he be disposed to utter anything?

5. Self-talk and saving face

One plausible scenario in which the messy shopper could utter 4 is one of self-directed speech, that is, self-talk. Self-talk seemingly violates certain rules about why and how we talk, since paradigmatically, talking involves two

31 Campbell 1994: 135.

32 Goldstein 2007: 59.

33 Bloom 2000: 122.

34 See Campbell 1994: 141–142 on the same experiment.

or more agents – taking up roles as either the speaker or the hearer – who communicate in order to exchange information. This marks a strong contrast with self-talk, which, oddly enough, involves agents addressing themselves to share information that they already have.

One way to make sense of it is to abandon the idea that we talk to share information and consider, for example, cases in which we utter sentences just to manifest our inner discourse. In these cases, talking is typically meant to aid silent reasoning and enable mental organization. The agent is not sharing information *per se*, but rather organizing the information she already possesses by verbalizing certain thoughts and helping herself identify inferential relations. This “platonic” understanding of self-talk³⁵ allows us to explain self-directed utterances exhibited in problem-solving or planning, such as 6 below, said by a self-talker planning her actions for the day:

6. I will stop by the gas station and then go to class.

It is not clear though how this platonic understanding applies to the very common cases of self-talk associated to self-conscious behavior and self-encouragement, like in 7. and 8. below, respectively³⁶.

7. What is wrong with me?
8. I can do this!

Most literature on self-talk looks into psychology and cognitive development to explain these last cases. The reason is that it has been assumed by some prominent psychologists, like Vygotsky (1981)³⁷, that self-talk plays a part in the internalization of social relationships. This assumption is supported by observations that overt self-talk though highly frequent in early years becomes gradually more restricted in adult life³⁸, as individuals approach higher levels of sociability and begin to keep their inner dialogues more and more to themselves.

However, the platonic and the developmental perspectives give us just one part of the explanation, specifically pertaining to the socially acceptable cases of self-talk. They do not address why self-talk is typically inappropriate from a sociological as well as sociolinguistic standpoint. Indications of how to handle these worries can be found in Goffman (1978). As he remarks,

35 We will call this understanding ‘platonic’ in reference to Plato’s mention of inner dialogues in a somewhat similar sense in the *Theaetetus*.

36 Extracted from Holmberg (2010). As discussed by Holmberg, this kind of self-talk may also involve second-person pronouns.

37 As mentioned by both Goffman (1978) and Geurts (2018). Geurts, in particular, remarks that a more developmentally-oriented approach – than the purely platonic one – has the additional benefit of accounting for the continuity between self-talk and social talk according to a unified view of communication that is not centered on the idea of informational exchange.

38 Vicente and Martinez Manrique 2011.

adult lonely self-talk is taboo both because it violates the sender-recipient presupposition in communication, and because talking is paradigmatically a *situational* activity that involves the gathering and interaction of and between people. He defines a social situation as the “physical surroundings in which people gather to interact”³⁹. For him, the most essential trait of a social situation is the “respectful aliveness” that one is expected to show *in* and *to* it. Hence, self-talk is a social impropriety because it reveals that the speaker is not really *in* the social situation but rather finds herself distracted by inward states.

Nonetheless, there will be cases in which self-talk in social gatherings is not only adequate, but expected⁴⁰. Take one example inspired by Goffman. Imagine a situation in which two friends are walking together deeply engaged in conversation. One of them sees something shinning on the floor, something that looks like a coin. She stops to pick it up but after a closer inspection realizes that it is just the metallic lid of a yogurt bottle. She utters 9. to herself, and then returns to the conversation.

9. I thought it was a coin.

In Goffman’s analysis, this self-talker utters 9 to re-establish herself in the eyes of her witness as a competent person – “not to be trifled with”. She is *saving face*. Her comment is not being addressed to herself in order to exchange information, since she already has that piece of information. What her utterance really aims at is to justify her behavior: in uttering it, the self-talker blocks any doubt her friend might come to have regarding her rationality. As Goffman puts it, “an understandable reaction to an understood event”⁴¹.

And we may exhibit a similar behavior in the absence of direct witnesses too. Consider another example from the same text:

An unaccompanied man – a single – is walking down the street past others. His general dress and manner have provided anyone who views him with evidence of his sobriety, innocent intent, suitable aliveness to the situation, and general social competency. His left foot strikes an obtruding piece of pavement and he stumbles. He instantly catches himself, rights himself more or less efficiently, and continues on. Theretofore, his competence at walking had been taken for granted by those who saw him, confirming their assessment of him in this connection. His tripping suddenly casts these imputations into doubt. Therefore, before he continues, he may well engage in some actions that have nothing to do with the laws of mechanics. The remedial

39 Goffman 1978: 790.

40 Perhaps the utterance of ‘I can’t believe it’, when one is informed of the death of a loved one is the best example. Here, self-involvement is considered reasonable and not a demonstration of lack of respectful aliveness in and to the situation (Goffman 1978).

41 Goffman 1978: 797.

work he performs is likely to be aimed at correcting the threat to his reputation, as well as his posture. (Goffman 1978:793)

Call this *the stumbling man example*. Imagine that, after tripping, the stumbling man briefly looks around with a shy smile, as he utters 10 to himself:

10. I almost fell.

Unlike self-talk of the self-directing type – as in 8, with the intention of encouragement – or manifestations of inner discourse in self-negotiating plans, as in 6, 10 reveals the agent's inclination to respond to the situation as a social situation even in the absence of an addressee⁴². That is, even if he is addressing his words to no one in particular, the stumbling man is likely to, in Goffman's words, "tell a story to the situation" by externalizing a certain narrative – which can even be non-verbal, such as examining the pavement or his footwear to see what caused him to stumble.

What is really distinctive in this example is the overt attempt to save face when confronted with a potential threat to one's reputation. Like 9, the utterance of 10 can be seen as an anticipatory move in case remedial action is necessary – even if talking to oneself is still a somewhat strange behavior. Goffman claims, for example, that: "when circumstances conspire to thrust us into a course of action whose appearance might raise questions about our moral character or self-respect we often prefer to be seen as self-talkers" (:796). So, we not only accept but expect something like 10 to be uttered by our clumsy character.

Some cases of save face talk will involve taboo words, like expletives, in the case of 'Damn!', or so-called response cries, like 'Oops!'⁴³. The use of these expressions in self-talk too may serve to evidence that the agent is watchfully tracking her own movements in the social situation – although here she presents her rationality in terms of accountability for mistakes, failures and equivocations rather than in terms of remedial justification. In a save face emission of 'Oops!', for example, the agent's purpose is usually to insulate that specific behavior as an exception to her generally competent conduct. By recognizing a specific episode as an equivocation that warrants a 'Oops!', the agent saves face by trying to restore her image as a competent and respectable counterpart who is very much *in* the situation. Something similar happens with 'Damn!'; the main differences being that an inner state of frustration is also conveyed and that the expression has offensive potential. One can imagine the tardy professor talking to his brother in his office and then interrupting the conversation to utter something like 11 to himself.

42 Goffman (1978) develops the notion of embedment which can be understood roughly as "the lifting of a form of interaction out of its natural place and, and its employment in a special way" (:790). It is because of embedment – and ritual constraints – that agents respond to situations *qua* social situations even in the absence of potential addressees.

43 Goffman 1978: 800–806.

11. Damn! The meeting starts *now*.

And then running in the direction of the department's boardroom. Here, we have a mix of coordination and self-talk. The tardy professor breaks the joint activity. But in exhibiting frustration with himself in front of his colleague, he also evidences awareness of the fact that he failed to provide what the situation required of him. The unconcealed recognition of his own momentary incompetence has the effect of restoring his good image as a rational agent "overall"⁴⁴.

Now, back to Goffman's stumbling man example, we believe that it resembles Perry's messy shopper story in many respects. Both characters are alone in potentially social (that is, public) environments – the street and a supermarket – and both come to realize that possible witnesses could perceive them as incompetent agents – clumsy, inattentive, laughable. Both seem disposed to utter 'I' – e.g., in 4 and 10 – to save face.

We also believe that the idea of a "save face speech act" helps us make sense of the dispositional assumption: John Perry, the shopper, will classify his action by using an indexical expression, namely 'I', not because he is constantly and indistinctly disposed to utter 'I' to predict and explain all of his thoughts and actions, but because making a mess at a Safeway store disposes him to say something for self-representational purposes. By uttering 4, he says to whoever happens to see the scene, that he was not making a mess on purpose, and that now that he sees that he is responsible for the trail of sugar on the floor, he will rearrange his sack.

6. Back to the dispositional solution

We have, at this point, a new interpretation of the messy shopper story, or, more precisely, of his disposition to utter a sentence such as 4. A first account was that he has a token of this sentence in his mind, as he realizes the he was responsible for the mess. However, even if this is true, having a sentence in the language of thought doesn't explain the disposition to utter anything.

Another way to see the story is that this indexical sentence classifies his actions because he is disposed to utter 'I am making a mess,' whether or not it corresponds to a sentence in Mentalese tokened in his mind. But it is not clear that one is disposed to say aloud just any thought that explains one's behavior. Imagine that I am driving and realize that I have to turn right at the next corner. There is no reason to think that, as I realize that, I utter, or have a disposition to utter 'I will turn right now,' even if the thought that I will turn right now explains my actions.

44 In the original example, the tardy professor is alone in his office, probably behind closed doors. He would be disposed to self-talk most likely for the purpose of mental organization, as in the example of 6, and not to save face.

We may have another take on indexicals as classificatory devices. One way to follow this path is to accept some sort of structural similarity between indexical roles in thought and indexical meanings in language, along the lines suggested by Robert Matthews for the attribution of propositional attitudes.⁴⁵ Maybe so, but we still lack an explanation, if there is indeed any, of the disposition that the messy shopper has to say ‘I am making a mess.’

There are cases in which an agent may utter a sentence describing what she is doing in order to coordinate her actions with what her partner is doing, or with what her partner expects of her. The messy shopper, however, is not engaged in any joint action. Nevertheless, we think that the messy shopper *is* disposed to utter 4 not because he realizes what he was doing, but because, on top of that, what he was doing threatened his image as a competent agent. Interestingly, the story strikes us as plausible because of something that Perry doesn’t say: the only bearded philosopher in a Safeway store west of the Mississippi is *disposed* to utter 4 because of his clumsiness and the consequent need to save his face.

7. Concluding remarks

Indexicals classify mental states in order to explain and predict actions. We need to distinguish what is believed from how it is believed, and “we use indexicals to disclose *how* we believe; singular propositions get at *what* we believe.”⁴⁶ However, we cannot classify the way an agent thinks of a situation by the sentence she is disposed to utter: the disposition to do something is not the disposition to utter a sentence describing what one is doing. Indeed, arguing that indexicals are needed in the explanation and prediction of an action, and, thereby, that the words ‘I’, ‘here’ and ‘now’ are essential, leads to a dilemma: there is no room for ‘I’, ‘here’ and ‘now’ in a classificatory view of indexicals, and the argument is supposed to establish, precisely, the ineliminability of indexicals.

We can avoid this dilemma and accept Perry’s solution by acknowledging that we are after the roles of indexical thoughts in our mental lives and that indexicals in language, inasmuch as they are part of the activity of explaining

45 “One of the principal tasks for a measurement-theoretic construal of propositional attitude attributions is to specify the empirical structure of the attitudes that has its image in the representational structure just described. As is the case with other sorts of measure predicates, competent users of propositional attitude predicates are typically unable to specify this structure. It will therefore have to be reconstructed inferentially, from a consideration of (i) the empirical evidence on which attitude attributions are based and the *explanatory and predictive tasks to which they are put*, and (ii) the structure of our natural language representations of the attitudes, on the assumption that the relevant properties and relations of propositional attitudes about which we reason surrogatively using those representations are respected <...> by those representations.” Matthews 2007: 173 (italics added)

46 Perry 2020: 17.

and predicting actions, classify such roles. It is not difficult to disentangle this view from the link between indexical beliefs and the utterance of sentences with indexicals via the disposition to utter such sentences. In particular, as we go towards a more instrumental view of indexicals as classificatory devices, we are less tempted to accept the dispositional assumption. Already in the 1979 paper, Perry views indexicals as classificatory devices, and the dispositional assumption is not essential to his argument.

However, in Perry's story, there is something more interesting and somehow unexpected, or at least, to the best of our knowledge, unacknowledged so far. The messy shopper is, after all, disposed to utter 'I am making a mess', not because of his realization and change in behaviour – he could, after all, have changed his course of actions without uttering a word. He is disposed to utter this sentence because of his clumsy behaviour which threatens his image as a competent agent. This is a case in which, according to Goffman, self-talk is accepted, if not expected. In order to show that, in spite of the mess, he is a competent agent, he says 'I am making mess' or otherwise shows that he had not realized that he had a torn sack of sugar in his cart. He shows it to whoever happens to see the scene. For their eyes only.

References

- Austin, J. (1979). *3rd Philosophical Papers*. Oxford: Clarendon Press.
- Bloom, P. (2000). *How Children Learn the Meaning of Words*. Cambridge, MIT Press.
- Campbell, J. (1994). *Past, Space, and Self*. Cambridge: MIT Press.
- Campbell, J. (2003). The Role of Demonstratives in Action-Explanation. In Eilan, N., & Roessler, J. (eds.), *Agency and Self-Awareness*, Oxford: OUP, pp. 150–164.
- Corazza, E. (2004). *Reflecting the Mind – Indexicality and Quasi-Indexicality*. Oxford: Oxford University Press.
- Crawford, S. (2008). Quantifiers and Propositional Attitudes: Quine revisited. *Synthese*, 160(1), pp. 75–96.
- Geurts, B. (2018). Making sense of self talk. *Review of philosophy and psychology*, 9(2), pp. 271–285.
- Gilbert, M. (2014). *Joint Commitment – How we make the Social World*. Oxford: Oxford University Press.
- Goffman, E. (1978). Response Cries. *Language*, 54(4), pp. 787–815.
- Goldstein, E. B. (2007). *Sensation and Perception*. Belmont: Wadsworth Publishing.
- Kaplan, D. (1968). Quantifying In. *Synthese*, 19(1), pp. 178–214.

- Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and other Indexicals. In Almog, J., Perry, J., & Wettstein, H., (eds.), *Themes From Kaplan*. Oxford: OUP, pp. 481–563.
- Korta, K., & Perry, J. (2011). *Critical Pragmatics*. Cambridge: Cambridge University Press.
- Matthews, R. J. (2007). *The Measure of Mind: Propositional Attitudes and Their Attribution*. Oxford: Clarendon Press.
- Perry, J. (1979). The Problem of the Essential Indexical. *Noûs*, 13(1), pp. 3–21.
- Perry, J. (1997). Indexicals and Demonstratives. *A Companion to the Philosophy of Language*, pp. 586–612.
- Perry, J. (2000). *The Problem of the Essential Indexical and Other Essays*, Expanded Edition. Stanford: CSLI Publications.
- Perry, J. (2001). *Reference and Reflexivity*. Stanford: CSLI Publications.
- Perry, J. (2007). Situating Semantics – A Response. In O’Rourke, M., & Washington, C. (eds.) *Situating Semantics – Essays on the Philosophy of John Perry*. Cambridge, Mass.: MIT Press, pp. 507–575.
- Perry, J. (2020). *Revisiting the Essential Indexical*. Stanford: CSLI Publications.
- Quine, W. V. (1956). Quantifiers and Propositional Attitudes. *Journal of Philosophy*, 53(5), pp. 177–187.
- Recanati, F. (2000). *Oratio Obliqua, Oratio Recta: An Essay on Metarepresentation*. Cambridge: MIT Press.
- Recanati, F. (2012). *Mental Files*. Oxford University Press.
- Vicente, A., & Martinez Manrique, F. (2011). Inner Speech: Nature and Functions. *Philosophy Compass*, 6, pp. 209–219.
- Vygotsky, L. (1981). The Genesis of Higher Mental Functions. In Wertsch, J. V. (ed.), *The concept of activity in soviet psychology*, pp. 144–188. Armonk: Sharpe. Originally published in Russian in 1960.



DEMYSTIFYING THE MYTH. PERRY: REVISITING THE ESSENTIAL INDEXICAL

We often feel inclined to express certain attitudes using sentences containing indexicals. These are, paradigmatically, cases of beliefs about oneself, which we normally express with sentences containing the first-person pronoun “I,” or about the present time, which we normally express with sentences containing the temporal adverb “now.” One could, of course, express a belief about oneself using a sentence containing a proper name, or about the present moment with a sentence containing a date. But there are, it seems, clear and important features that make “indexically expressible beliefs” different from other sorts of beliefs. Or so did Perry, among others, claimed in his 1977 and 1979 papers (and has kept claiming ever since).

Perry was not the first, and not the only one, to realize that there is something philosophically relevant about indexicals and indexically expressible beliefs. Before him, Castañeda (1966, 1967, 1968) and Prior (1959, 1968) made observations in a similar direction; and, roughly at the same time, Chilshom (1981) and, most famously, Lewis (1979).

Since then, many authors have commented on the role of indexicals and of indexically expressible beliefs. Some have defended Perry’s proposal, or parts of it, at least. I call them, somehow dramatically, the “apologists.” Many others, however, have either disregarded the whole issue or have not accepted Perry’s explanation. Those who realize the importance of the essential indexical, but disagree with Perry’s account, I call “skeptics.” Those who disregard the discussion altogether and consider, for instance, the essential indexical a “myth,” I call “negationists.”

In his book *Revisiting the Essential Indexical* (2019) Perry mostly deals with negationists and, in particular, with Cappelen and Dever, who defend that:

...there is no such thing as essential indexicality, irreducibly *de se* attitudes, or self-locating attitudes. Our goal is not to show that we think these phenomena – that they should be explained in ways different from how, e.g. Lewis and Perry explained them. Our goal is to show that the entire topic is an illusion – there is nothing there (Cappelen and Dever, 2013: 3)

Perry offers a clear, comprehensive and devastating answer to Cappelen and Dever's claims in *The Inessential Indexical* (2013). In this paper my goal is to emphasize the importance of this book and to insist on clarifying some central issues that have been largely ignored by many participants in the debate. In the first part of this paper I shortly comment on Perry's response to Cappelen and Dever, showing how they also apply to other negationists. In the second part, I discuss Arthur Prior's example, and Perry's take on Lewis.

Some comments on *Revisiting the Essential Indexical*

Anyone with even a slight interest in the philosophy of language has surely heard of a messy shopper (Perry himself) spilling sugar on a supermarket. Most would also be familiar with Heimson, who believed he was Hume; with the tardy professor, late for a meeting; with the lost hiker, at Gilmore Lake; and with amnesiac Lingens, lost in the Stanford library. These examples, presented in Perry's 1977 and 1979 papers have proven to be very catchy. During the last decades, Perry has enriched and elaborated his views substantially, but the examples used in his earlier papers have become part of the philosophical *cannon*, so to speak. They have been quoted, used, retold and reinterpreted countless times. Often, however, the lessons that have been extracted from them, and the explanations given of them, have little in common with Perry's (both in those early papers or in later elaborations).

I take Cappelen and Dever's view, as expressed in their 2013 book, to be an extreme case of misinterpretation of Perry's examples and of Perry's proposal. Cappelen and Dever reject the significance of the examples themselves, and of what they display. They call themselves "skeptics," I call them "negationists."

Negationists distort not only the explanation of the issues, but the issues themselves. Cappelen and Dever do so, mostly, to argue against what they call "perspectivism:" the claim that there are certain beliefs that include the agent and their situation in time in their contents. All beliefs, they claim, are about the "objective" world, that is, about the world itself.

They argue that admitting that indexicals generate or illustrate some particular puzzles to the explanation of belief content, causal role and cognitive significance, is adamant to a renunciation of objectivity. But of course, none of the examples mentioned above, and certainly not Perry's explanation of them, entail (nor presuppose) that indexically expressible beliefs are not about the world *in itself*. To use Falk's expression, Cappelen and Dever are "whipping a straw man." (2015, 427). In my opinion, this description captures pretty well Cappelen and Dever's whole enterprise in their book. And I think Perry shows why and how this is so.

Perry answers to Cappelen and Dever's harsh criticism in a clear, direct, structured and very convincing manner. His arguments are withering. It is difficult to imagine how Cappelen and Dever could reasonably respond to this (they haven't tried, as far as I know). Perry doesn't answer to all the claims and criticisms of Cappelen and Dever. That would have been not only pointless but also rather boring (to read and, I suspect, to write). Rather, he focuses on two or three main confusions and answers them.

Cappelen and Dever present, discuss, and reject six kinds of "alleged explanatory roles" of indexicals. Perry focuses most of his discussion on the first two. They are directly attributed to him. They are also the fundamental ones, I think, for a discussion on the role of indexicals in thought and action.

Agency: [...] the idea that indexicality (and "the *de se*" in particular) plays an essential role in explaining and rationalizing action.

Opacity: [...] the question of whether the presence of indexicals in (apparently) opaque contexts raises questions that are fundamentally and interestingly different from general issues about opacity. (2013, 14–15. Boldface in the original)

Cappelen and Dever assume that Perry defends **Agency**, and they argue that indexicals do not raise questions different in any significant way from general issues about opacity. As Perry clearly explains, however, he never claimed that "*indexicality* plays an essential role in explaining and rationalizing action, but that in particular cases *indexicals* did" (Perry, 2019, 18). Cappelen and Dever take indexicals (or indexicality) as the *explanans*. Perry only claimed that they were part of the *explanandum*:

[W]hat did the explaining was the distinction between the how and the what, between belief states and what is believed... Using indexicals exhibits a pattern —the distinction between *how* one does something and *what* one does in that way— that certainly plays an essential role in explaining and rationalizing action, recognized by (almost) any theory of action. (Perry, 2019, 18)

It is clear from Perry's earlier texts that he took indexicals to be part of the sentences that give you the phenomena to be explained. The difference between "John Perry is making a mess" and "I am making a mess" is the phenomenon to be explained. Indexicals are part of the phenomenon, part of the puzzle, not its explanation. I will come back to this distinction in the next section, when discussing ways of classifying beliefs.

Cappelen and Dever's **Opacity** is, as Perry says "more mystifying," but it is not entirely original. The claim is also made —with some differences— by two other notable "negationists": Millikan (1990) and Magidor (2015). To this, Perry answers:

The Confusion: Cappelen and Dever confuse opacity with cognitive significance. This is the leitmotif of their book and the basis of most of their criticisms. But, as they might put it, there is nothing there. (2019, 19)

Perry discusses at length this confusion. In an attempt to make sense of it, he tries a very “charitable” reading of their claims, substituting “opacity” for “cognitive significance.” This is, indeed, a very charitable reading. **Opacity** is, perhaps, the central claim of Cappelen and Dever’s book. They present “Fregean counterparts” for most of Perry’s, Lewis’s and Prior’s examples. Still, even changing “opacity” for “cognitive significance” their main claim does not hold. There is something particular about indexicals, something that distinguishes them from proper names and other referring terms.

I will not give many details about Cappelen and Dever’s arguments, or about Perry’s answer to them. That would take too long and, besides, Perry’s discussion is concise enough and very clear. For those who haven’t read the book yet, however, I write below Cappelen and Dever’s tenets on opacity — their “boldface claims”— and Perry’s reconstruction in terms of cognitive significance. This short presentation will be enough, I think, to illustrate the confusion mentioned above, and why it remains so under a charitable interpretation. Here is what Cappelen and Dever say:

Here is one way to summarize Perry’s claim:

Indexical Opacity. There’s a set of indexicals, *I-SET*, that cannot be substituted *salva veritate* in action-explanation contexts by any other expressions.

It should be clear that this is an instantiation of the more general thesis:

Generic Opacity. Co-referring expressions cannot be substituted *salva veritate* in action-explanation contexts. (2013, 33)

They proceed by offering a “Fregean counterpart” of the messy shopper example, substituting “Clark Kent” and “Superman” for “John Perry” and “I.” They conclude that “seeing Indexical Opacity as an instance of Generic Opacity suggests that there’s nothing deeply central about indexicals here” (2013, 33).

Perry finds this way of summarizing his claim “quite puzzling.” I share his sense of puzzlement. None of the examples Perry gives involve or suggest that substitution of indexicals changes the truth-value of sentences. Also, on Perry’s view, like in Kaplan’s, attitude reports are *not* opaque.

Substituting “John Perry” for “I” in “I am making a mess” can be done *salva veritate*. It might be that the proposition expressed changes —depending on your view on propositions—, but certainly not the truth-value. What seems to change, in “action-explanation contexts,” is the cognitive significance. So, this is Perry’s reformulation:

Indexicals and Cognitive Significance. Substituting indexicals with co-referential expressions may change the cognitive significance of the sentences in which they occur, including their explanatory force.

Referring Expressions and Cognitive Significance. Substituting any referring expression with co-referential expressions may change the cognitive significance of the sentences in which they occur, including their explanatory force. (2019, 33–4)

As I mentioned before, **Opacity** might be “mystifying” but it is not entirely original. Ruth Millikan defended something quite similar, in her paper “The myth of the essential indexical” (1990), and so did Magidor in “The myth of the *De Se*” (2015). There are some differences of detail among the three, of course, but they all think that the philosophical questions that indexicals raise for action and thought are illusions, “myths.” To sustain their views, and ignoring questions of detail, they argue that the issues Perry and Lewis discuss concerning indexicals are just instances of Fregean puzzles. These puzzles involved proper names, and not indexicals. So, there is nothing particularly problematic or “deep” or “philosophically relevant” about indexicals or, as they all say, about *de se* beliefs. More importantly, taken as instances of Fregean puzzles, negationists defend that a proper treatment of indexicals, and a proper explanation of Perry’s and Lewis’s examples do not require any modification to the traditional notion of propositional attitudes (the “doctrine of propositions”). It is not my aim to refute, or even discuss negationists’s arguments. But I think it is worth mentioning several points.

First, most, if not all, focus on *de se* attitudes, which they wrongly attribute to Perry. Some, like Magidor, conclude that there is nothing philosophically interesting about Perry’s examples, and about indexically expressible beliefs. But she reaches this conclusion after discussing and rejecting Lewis’s account of Perry’s examples.

She only discusses Perry proposal on a footnote (footnote 3, 275). And she does so to claim that

... the view he [Perry] opts for arguably has the resources for addressing the Frege’s puzzle more generally... but Perry should at least be considered to be a defender of the myth in so far as *he* construes the issue as a special challenge posed by *de se* beliefs. Moreover, Perry clearly supports Special Attitudes. Fn 3, 275. Her italics).

She concludes,

I maintain that the category of *de se* attitudes (if there is one) does not play any important role in the semantics of attitude reports or require any special amendment of our general account of propositional attitudes. The myth of the *de se* remains just that. (Magidor, 2015, 272)

But, as I explain in the next section, Perry does not defend that we need a special attitude to account for indexically expressible beliefs. What he proposes is that we talk of belief states and recognize two ways to classify them: according to what is believed and to the way in which one does it. Lewis, who coined and uses the expression “*de se*”, understands *de se* beliefs as the self-attribution of a property. So, he does seem to be claiming that we need a special attitude: self-attribution. It appears that Magidor is attributing to Perry Lewis’s claims and definitions. Even if her arguments against *de se* were correct, they would not prove much about Perry’s explanation of indexically expressible beliefs and the challenges they pose. The *de se* might well be a myth, as she claims, and yet Perry’s puzzles and explanation remain real and philosophically relevant.

Second, Fregean counterparts are only possible for what Perry calls “Type B” cases: where two people, or the same person in different spatiotemporal locations, do the same thing in different ways. Type B examples are the messy shopper, the lost hiker or the tardy professor. “Type A” cases are those where two people, or the same person in different spatiotemporal locations, do different things in the same way. An example of type A is Heimson and Hume.

Negationists only consider Type B cases. Cappelen and Dever claim that Type A ones are simply irrelevant. Considering that Type A examples are meant to illustrate cases in which *what* is believed is different but the *how* is the same, it is no wonder that Cappelen and Dever think that Perry defends a very implausible and extreme form of perspectivism, according to which indexically expressible beliefs are “not objective.” Blurring the difference between the content of beliefs and ways of believing them is, in my opinion, what makes them think that the *how* is part of the *what*: that the ways of believing are part of what it is believed.

Classifying beliefs

Castañeda talked about the indexicals “I” and “now” being “essential” because they cannot be defined in terms of other referring expressions. Perry uses the term “essential” in a similar way, to indicate that there is something, some information, that can only be expressed by an indexical. We use indexicals to refer to objects —oneself and the present time, respectively, in the case of “I” and “now”— and they convey certain information about these objects not conveyed by proper names, descriptions or dates.

When they are part of sentences that express what one believes, indexicals play an “essential” role in exactly that sense: in conveying certain information about the object of the belief. This information is essential to get at the content of our beliefs —what we believe— but also to understand the causal role beliefs have —for subsequent actions, mostly— and their

cognitive significance —how it is that one might believe A and not believe B, even though A and B seem to have the same content and truth-value.

Roughly speaking, the basic idea is that there is a difference between beliefs that are about how the world is “in itself” and beliefs about the agent’s situation in the world. Perry calls the latter “locating beliefs” (1977) or “self-locating beliefs” (1979): “one’s beliefs about where one is, when one is, and who one is. Such beliefs seem essentially indexical” (Perry, 1979: 29).

How to classify beliefs is a matter of controversy. Following basic intuitions from folk psychology, it would seem that there is a difference in what one believes and how one believes it and, consequently, how one would normally express it. These are, if I am not wrong, the intuitions Perry uses and the distinction he exploits. To put it in different, and perhaps more accurate terms: it is one thing to believe the same thing and another quite different to be in the same belief state.

Consider Perry’s tardy professor (1979, 29) and his two beliefs, expressible as:

The department meeting begins at noon.

The department meeting begins now

Assuming that noon is the time of belief state —and, hence, the time referred to by “now” in (2)— in one sense, it seems that the professor would express one and the same thing with (1) and (2): that the meeting begins at 12:00 of the relevant day. What he believes seems to be the same on both cases. In another sense, however, (1) and (2) clearly seem to express different beliefs in at least three ways. First, the sentences used to express them are different: one includes an indexical and the other a noun. Second, the causal role of each belief seems to be quite different: (1) might make him write down “meeting” besides the number 12 in his calendar; (2), in contrast, should make him leave the office and go to the meeting (Perry’s professor is tardy, but responsible). Third, their cognitive significance is different: the professor might, and actually does at first, believe (1), but not (2).

One should be able to say that two people believe alike if they believe the same content or proposition, even if they are at different times or places, or even in different belief states. Similarly, one should be able to say that if two people believe, at noon, (1) or (2), they agree; even if they don’t realize it. This accounts for the intuition that the tardy professor believed the *same thing* in (1) and (2) (again, assuming the belief expressible by (2) happens at noon). Cases like this one —like the messy shopper and the lost hiker—are what Perry calls Type B cases: those where the same thing is done, in different ways.

Perry also discusses the opposite type of cases, Type A: those in which different things are done in the same way. To use another classic example by Perry (1977, 16), if Heimson and Hume both believe what they could express as:

I am Hume
I wrote the *Treatise*

they would both be in the same belief state. But what they believe, the proposition they believe is quite different. Hume believes of himself that is Hume and wrote the *Treatise*, which is true, and Heimson believes of himself that is Hume and wrote the *Treatise*, which is false.

Type B cases, with examples like the tardy professor with regards to “now” and the messy shopper with regards to “I” were specially designed to deal with the role beliefs play in explaining action. Even though, as we said, the content of the beliefs of the tardy professor is the same in (1) and in (2), the causal role they play in his subsequent actions is definitely different, and so is their cognitive significance. The tardy professor is in belief state (1) all along, but he is not in (2) until it is late for him to get to the meeting on time. His *coming to believe* (2) —that is, coming to the state belief (2)— causes him to stand up and run to the department’s meeting room.

Type A cases, with examples like Heimson and Hume’s or Lewis’s two gods scenario are useful to explain “sameness of belief state,” or the role the proposition believed plays in accounting for what we believe, and in dealing with agreements and disagreements about what we believe.

Perry’s explanation for all these cases is similar. It involves the already mentioned difference between the *thing* believed and being in a certain *belief state*, and some modifications to what Perry calls “the doctrine of propositions,” which involves the following three tenets:

- Beliefs (and other cognitive attitudes) *consist in* relations to the propositions referred to by the ‘that’-clauses of attitude reports of the form “X believes that S.” S is the *embedded sentence*.
- The truth-values of propositions do not depend on who asserts them or believes them, or when.
- If X believes the proposition that S, that belief will lead X to regard S as true. That is, the proposition not only captures the truth-conditions of S, but also its cognitive content or cognitive significance, the beliefs that lead one to regard it as true. (Perry 2019, 5)

Perry argues for modifying the first tenet, keeping the second and rejecting the third.

The traditional notion of proposition, and the received view on propositional attitudes cannot accommodate indexically expressible beliefs. Both Fregean and singular propositions struggle with them. The received view of propositional attitudes, captured in the three claims above, cannot accommodate these cases, and the accompanying intuitions.

Perry and Lewis agree in this point. Lewis also defended that beliefs need to be classified in terms of roles or, as he called them, properties.

Propositions, Fregean or singular, are not enough. So it might be correct to talk about a “Perry-Lewis view,” if it is used to refer to the claim that we need to give up the doctrine of propositions. Or parts of it, at least. But Perry and Lewis do not agree on what needs to be rejected or kept from it and, most importantly, what should we substitute for it. I’ll say something more about this point of agreement, focusing on Perry’s account of indexicals. Then, I will shortly contrast his views with those of Prior and Lewis, following what he says about them in the book.

Frege, “I”, and the order of quantifiers

Some of the features of the examples above are similar to Frege’s puzzles involving proper names. But, contrary to what negationists claim, some are clearly distinct. Frege was aware that indexicals raise particular problems, and, in his later works, he tried to accommodate them on his account. In his 1977 paper “Frege on demonstratives,” Perry argued that Frege failed in that attempt, and that indexicals (and demonstratives) demanded some amendments to Frege’s view. In his 1979 paper, he argued that alternative accounts on meaning and propositions—which understood them as singular—did not fare better in dealing with them.

Just like with proper names, we do not need to know much to be able to refer by using “I” or any other indexical, but we need to know something: the indexical’s role. “Role” is the term used by Perry to designate the functions that “take different people at different times who are in the same belief state to different things believed” (2019,12). Consider Perry’s famous example, “the messy shopper.” Perry is following a trail of sugar on a supermarket floor. After a while, he realizes that the sugar is falling from a torn sack in his cart: he comes to believe that he is the one making a mess. Imagine he uses one of the following to express —aloud or to himself—his new belief:

John Perry is making a mess

I am making a mess

John Perry might have used (5) or (6) to express what he believed when he realized that *he* was making a mess. “John Perry” and “I” refer to the same person in (5) and (6), but if we substitute “I” for “John Perry,” the cognitive effects (and motivations) would radically differ. (6) would be perfectly normal if John Perry wants to express what he is doing at a particular moment in time. On Perry’s example, it is a natural way of expressing what he came to realize, i.e. that he was the one making a mess. He can now stop chasing the mysterious messy shopper, and start fixing the sugar sack in his own cart. (5) would be a true but weird and pompous way of saying what he realized and came to believe. On most contexts, believing that “John Perry is making a mess” would make him stop and fix the situation. This is because he knows

that he is John Perry. But (5) would not make him check his cart if he doesn't know he is John Perry, or if he thinks there is another person called "John Perry" at the store.

Besides, upon hearing (6), anyone would understand that John Perry is expressing what he realized of believes he was doing; but only those who know that his name is "John Perry," and that for some reason he is referring to himself in the third person, would directly understand this upon hearing (5). At the very least, they would be prompted to ask something like, "what do you mean, that you are the one making a mess?"

The role of the indexical gives the conditions to identify the referent, by determining how this referent is presented. The role of "I" indicates that the referent is the speaker of the utterance, or the agent of the belief, containing "I." The referent—the speaker/agent—is thus presented in a certain way: in a first-person way. So, even if you do not know anything about John Perry, you will know that he is talking about his actions and his beliefs upon hearing him utter (6).

The role of an indexical however is not, and cannot be, a Fregean sense, because it gives us a different object on each occasion of use. Frege's solution to Frege's puzzles doesn't seem to work when indexicals are involved. And it doesn't because indexicals present different characteristics than names, and thus require a special treatment. Think of (5) and (6). Frege claimed that sentences express Thoughts and refer to truth-values. A Thought, for Frege, is a complete, objective and invariant entity. It is what we objectively understand when we understand a sentence, independently of our own particular thoughts or impressions about it. This, of course, is what guarantees communication. The Thought expressed by (5) is something like "John Perry is making a mess at t." The Thought expressed by (6), when John Perry is the speaker and assuming it is uttered at the same moment in time, would also be "John Perry is making a mess at t."

But this cannot be the end of the story. The Thought expressed by (5) would remain the same regardless of who utters it. The Thought expressed by (6) changes from speaker to speaker. How can they be the same Thought then? And how can the Thought expressed depend on who utters it? Assuming there is a sense associated to the "I," this sense would need to be different for each speaker, since, for each speaker, a different Thought is expressed upon uttering (5) above. But then, if these Thoughts are unique for each speaker, how can we communicate them?

Frege's writing on this issue are very few and came quite late in his career, mostly in his paper "Thought" (1918–1919). Consider his most often quoted paragraph on "I,"

Now everyone is presented to himself in a special and primitive way, in which he is presented to no one else. So, when Dr Lauben has the

thought that he was wounded, he will probably be basing it on this primitive way in which he is presented to himself. And only Dr Lauben himself can grasp thoughts specified in this way. But now he may want to communicate with others. He cannot communicate a thought he alone can grasp. Therefore, if he now says “I was wounded,” he must use “I” in a sense which can be grasped by others, perhaps in the sense of “he who is speaking to you at this moment;” by doing this he makes the conditions accompanying his utterance serve towards the expression of a thought (Frege. 1918–9: 333. Note omitted)

The idea that each of us is presented to ourselves in a special way is of course old. At first sight, Frege seems to be appealing to the common idea that we are aware of ourselves in a first-person way; i.e. that there is an important difference between first-person and third-person knowledge. Nothing particularly problematic here. Unless we attempt to accommodate it within Frege’s account of senses as complete, objective Thoughts.

Actually, it is not entirely clear what Frege is saying in the above paragraph. The first sentence, in particular, admits two readings, as Perry (2019: 8–9, 52–3) points out. Frege might be saying that there is a special and primitive way in which each person is presented to oneself, and that this is different to the way she is presented to others. This is a pretty uncontroversial claim, and it is difficult to see why Frege would find it troublesome.

But there is a second reading, according to which each person has one special and primitive way of presenting to oneself, which only that person can grasp and which she cannot therefore communicate. This particular and primitive way of presenting to oneself is a primitive way of knowing about oneself, and it is expressed with particular and primitive Thoughts: Thoughts that only each of us can grasp. In other words, a person cannot communicate the Thoughts grasped in this particular and primitive way.

The difference between the two ways of understanding Frege involves a change the order of the quantifiers. On the first reading, the one favored by Perry to explain self-knowledge, what Frege would be saying is that,

There is a particular and primitive way in which every person is presented to himself, and no one else (2019, 9)

On the second reading, which seems to be the one Frege intended, what he is saying is that,

For each person, there is a particular and primitive way in which he is presented to himself and no one else... and no one else can grasp Thoughts determined this way (2019, 8–9)

On this second reading, if Dr. Lauben utters, “I am wounded,” he would be expressing a Thought only he can grasp and one that, therefore, he cannot

communicate. That Thought includes the sense of “I,” which includes the primitive and particular way or presenting himself and which determines Dr. Lauben as the referent. Since only he can *access* himself in that particular way, only he can grasp the sense associated with “I,” and the Thought expressed by the sentence “I am wounded” (when uttered by him, of course). In Frege’s terms: “only Dr. Lauben himself can grasp thoughts specified in this way... He cannot communicate a thought he alone can grasp.” This reading is, clearly, much more controversial.

If this is right, Frege’s claim in this paragraph entails that each of us has their own particular mode of presenting oneself to our self, a mode that is particular to each of us, only accessible by each one and, in that sense, incommunicable. A person “cannot communicate a thought he alone can grasp.” After all, communicating it would be to make it accessible for others. In Perry’s terms, “What is needed is a primitive aspect of me, which is not simply one that only I am aware of myself as having, but that I alone have” (1977: 15).

But then, we would need to accept that some Thoughts are not objective and graspable to all. That some senses are of limited accessibility. Considering the efforts Frege put in his earlier publications in specifying the role Thoughts play in communication and their objective nature, making them different from subjective ideas and thoughts, it certainly seems an odd admission on his part. This, roughly stated, is the conclusion Perry reached in his 1977 paper, and the problem he tried to solve (in that and in the 1979 paper).

Frege’s puzzles about proper names raise many issues, but, contrary to what negationists like Cappelen and Dever, Millikan or Magidor claim, they are not the same as the puzzles raised by indexicals. The solution Frege gives to the puzzles regarding proper names becomes insufficient when it comes to indexicals. Something needs to be added, or changed.

An important point of Perry’s solution is reversing the order of the quantifiers, as shown above. With this, we avoid non-communicable and non-objective belief contents. Many accounts of indexically expressible beliefs ignore the possibility of this second reading. And many reconstructions of what Perry and Lewis said about them take them to be adopting Frege’s order of quantifiers; with all the consequences this implies. Thus, for instance, García-Carpintero takes Perry to be committed to the “non shareability” or “non objectivity” of first-person thoughts *de se* (2017), and Holton takes Lewis’s *de se* beliefs to be “non communicable” (2015). They are both wrong, I believe.

Thoughts about the present moment—or the present day—fare no better on Frege’s account. And they fare no better on an account of indexically expressible beliefs as non-shareable or non-absolute. Quite the contrary.

Perry, Prior, and the order of quantifiers

Consider Arthur Prior's famous paragraph, which includes utterances and a different cognitive attitude: relief.

[H]alf the time I personally have forgotten what the date *is* and have to look it up or ask somebody when I need it for writing cheques, etc.; yet even in this perpetual dateless haze one somehow communicates, one makes oneself understood, and with time references too. One says, e.g. "Thank goodness that's over!" [...] says something which it is impossible that any use of a tenseless copula with a date should convey. It certainly doesn't mean the same as, e.g. "Thank goodness the date of the conclusion of that thing is Friday, June 15, 1954," even if it be said then. (Nor, for that matter, does it mean "Thank goodness the conclusion of that thing is contemporaneous with this utterance." Why would anyone thank goodness for that?). (Prior 1959: 17)

Suppose Arthur is leaving the dentist's office after undergoing a painful root canal. What is it that makes him exclaim (7), and not (8) or (9) in these circumstances?

Thank goodness that episode is/be over as of now.

Thank goodness that episode is/be over as of 2 p.m. June 15, 1954.

Thank goodness that episode is/be over as of the time of this utterance.

The short answer is that these utterances, by the same person at the same time, would have different cognitive significance. (7) seems to express relief on its own. (8) would require some supplementation to express relief (i.e. today is Friday, June 15, 1954). (9) is a very odd sentence to utter coming out of the dentist—or in any circumstance that does not include a philosophy seminar.

So, in a sense, the three utterances *mean* something different. They convey or display different information: one might believe (7) and not believe (8)—and vice versa—and the causal role they play is also different. If Arthur is talking to a good friend, (7) should cause the friend being happy for him; (8) would only manage that if the friend knows the time and date of the utterance—and knows that Arthur knows this, and somehow understands why he is using (8) rather than (7).

But, of course, the content expressed by the utterances, and the content of Arthur's relief, what he is relieved about, is the same: that the root canal is over as of the time of the utterance (which happens to be 2 p.m. June 15, 1954).

So, Prior's example is very similar to Perry's cases. And it is clearly designed to show how sentences containing indexicals are adequate to express cognitive attitudes such as relief, and sentences containing dates are not.

Prior's paper, however, has been often interpreted as a defense of the existence of A-properties of time, and their primacy over B-properties. That is, as a defense of an ontological claim, rather than a discussion about indexicals and indexically expressible attitudes. Cappelen and Dever, for instance, consider an ontological interpretation described, but not endorsed, by Ted Sider (on an online document, no longer available). According to them "the argument, if it succeeds provides evidence that attitudes like relief don't attach to B-facts, but that's not even the beginning of an argument for a distinctive kind of opacity arising in connection with indexicals" (2013, 67). They do not believe, of course, that there is anything philosophically interesting in the differences between (7)-(9), other than a question of opacity. Here too, they offer Fregean counterparts (Perry discusses these counterparts, and Prior's example, on chapter 9 of his book). They seem to believe, however, that Prior's argument, contrary to Perry's, might be interesting *as* and ontological one.

I do not think this is correct. First, because I do not think Prior's argument should be seen as an ontological defense of A-properties (de Ponte and Korta, 2017). Prior is a well-known defender of presentism, and he does defend an A-theory of time on many of his writings. This might explain, in part, why his example is often seen as having ontological consequences. But there is nothing in his paper to indicate that he is making an ontological claim. In any case, even if it were to be read like that, I cannot see how it would prove that relief is not "attached" —whatever that means—to B-facts. Cappelen and Dever do not explain how this is so.

More importantly for us here, however, is that Prior's interest in opacity is nowhere to be seen in the paragraph quoted or in the paper. Prior does not say that the truth-value of (7) and (9), when uttered at the same time as the example suggest, is different. According to Cappelen and Dever:

Suppose now = time *t*. Then (1*) – "I am relieved that the event is over now"– and (2*) – "I am relieved that the event is over at time *t*"– get different truth values, because "now" and "*t*" aren't intersubstitutable. (2013, 67).

But this is far from clear. They could have different truth values. But it is plausible to think, as defenders of singular propositions do, that (7)-(9) are different ways of reporting the *same* thing. Granted, (2*) would be a true, but misleading and quite weird report. But not false. If this is so, they would have the same truth-value. What Perry calls a Type A case: different ways to do the same thing.

Once again, as Perry claims, Cappelen and Dever confuse opacity with cognitive significance; and they fail to see the particularities of indexicals with regards to the latter. Negationists in general, and Cappelen and Dever in particular, fail to see the philosophical interest of the paper, because they fail to understand the issue it deals with: how is it that two utterances that seem

to express the same proposition —because they are utterances of sentences with co-referential terms— are not both appropriate in one circumstance? Why is (7) the right choice, and not (8) or (9), if the three seem to *say the same thing*?

Prior's example, and Prior's views on indexicals, anticipate many of the claims made by Perry and Lewis twenty years later. Actually, Prior's views on the "now" was very much influenced by Castañeda. Another point of contact with Perry. Just as it happens with Perry's examples the "order of the quantifiers" is key to see Prior's example as requiring additions to our ontology or, rather, as additions to the ways we classify cognitive episodes—relief—and utterances.

It seems obvious that there is a particular way in which we are all presented with whatever is happening at a certain time, at *that* time. That is, that there is a special way in which present events are presented to us, quite different to the ways future or past events are presented. At a moment of time *t* we have a particular access to events, or things that happen at *t*: we can perceive them. We cannot perceive events that haven't happened yet, or events that have already happened.

But, of course, that doesn't mean that we cannot communicate *what* we perceive, or *what* we believe or think about present events at other times. And it doesn't mean that we will not have access to them at future moments of times. This would only be the case if we apply Frege's preferred order of quantifiers and accept something like: that there is a particular and primitive way in which present events, and only present events, are presented to the agent... and from no other moment of time can the Thoughts determined this way be grasped.

So, just as, on Frege's preferred order of quantifiers, only I can grasp the content of a thought or a belief about myself. Because only I can access me *as myself*. Only when at moment *t* we can grasp the content of a thought or a belief about *t*. Because only at that moment we can access to that moment, and the events happening then, *as the present*.

This way of looking at things quite naturally leads to the claim that there is something particular about the present moment; some property or feature we can only capture when we are at *that* moment, and that we cannot revive before or afterwards. This special property or feature is an A-property: the event we are perceiving has the property of being present, and, after the moment passes, it will lose it and get the property of being past. Time flows, and moments in time gain and lose properties constantly. That particular and primitive way of getting to know present events is, basically, the fact that we can only perceive the property of "presentness," when we are present. After the moment passes, the that property is lost. The content of a believe about a present moment *t* cannot be replicated when the *t* is future, or past.

This way of seeing things might be, as we said, the natural conclusion if we follow Frege's preferred order of quantifiers. But it is certainly not the natural conclusion if we follow Perry's preferred order of quantifiers: There is a particular and primitive way in which the events happening at each moment are presented at *that* moment, and not at any other moment.

It seems that what we require here are different ways of knowing or accessing events that happen at different moments of time. So, there is a primitive way in which we can know what is happening now: perception. We cannot perceive past or future events; we can remember or anticipate them. Not because these events gain or lose properties, but because of our temporal location with regards to them. The ontological conclusion does not follow. Or, rather, it does not follow as *naturally* as before.

Arthur is thankful, at time *t*, about the conclusion of an event by *t*. He is thankful, presumably, because it is in the past; that is, because he cannot perceive that painful event anymore. Whether or not that event, or the moment of time in which it occurs, have or doesn't have a certain property seems quite irrelevant. Arthur is not thankful because his root canal has the property of "being past;" he is thankful that he cannot perceive it anymore. Further, to account for the difference among (7)-(9), the nature and properties of *t* don't seem very relevant. That is, whether or not *t* is a static moment of time in a series of ordered moments of time (B-series); or a moment of time that has the property of being present at the time of (7), had the property of being future before (7) and has the property of being past after (7) (A-series).

Perry, Lewis, and some conclusions

Perry's approach to indexicals gives us the necessary tools to deal with them without complicating our ontology *unnecessarily*. This ontological "simplicity" entails, however, a certain complexity in the way we understand and classify beliefs and other mental states. This complexity is what Lewis tried to avoid with his proposal. But, by doing so, he complicated ontology substantially.

Perry and Lewis claim that their views are, or could be, compatible. Lewis acknowledges this in his 1979 paper. Perry proposes a possible interpretation of Lewis compatible with his own, on chapter 12 of *Revisiting the Essential Indexical*. Lewis and Perry agree that sentences containing indexicals, and indexically expressible beliefs, generate problems for the doctrine of propositions. They disagree on the extent of these problems and on the way to solve them. Lewis considers Heimson and Hume's example:

The second problem arises when we ask why Heimson is wrong. He believes he is Hume. Hume believed that too. Hume was right. If Hume believed he was Hume by believing a proposition, that proposition was

true. Heimson believes just what Hume did. But Hume and Heimson are worldmates. Any proposition true for Hume is likewise true for Heimson. So Heimson, like Hume, believes he is Hume by believing a true proposition. So he's right. But he's not right. He's wrong, because he believes he's Hume and he isn't.

There are two ways out. (1) Heimson does not, after all, believe what Hume did. Or (2) Heimson does believe what Hume did, but Heimson believes falsely what Hume believed truly. (1979, 525)

Perry's way out is the first: Heimson and Hume believe different things, but they are both in the same belief state, which they would naturally express with (5) "I am Hume." By being in that belief state, the content of Heimson belief is that "Heimson is Hume," which is false, and the content of Hume's is that "Hume is Hume," which is true.

Lewis disagrees. He acknowledges that there "is *some* sense that Heimson does not believe what Hume did." But claims that there must be a "central an important sense" in which they "believe alike." He rejects Perry's explanation:

Heimson may have got his head into perfect match with Hume's in every way that is at all relevant to what he believes. If nevertheless Heimson and Hume do not believe alike, then *beliefs ain't in the head!* They depend partly on something else, so that if your head is in a certain state and you're Hume you believe one thing, but if your head is in that state and you're Heimson you believe something else. Not good. The main purpose of assigning objects to attitudes is, I take it, to characterize states of the head; to specify their causal roles with respect to behavior, stimuli, and one another. If the assignment of objects depends partly on something besides the state of the head, it will not serve this purpose. The states it characterizes will not be the occupant of the causal roles." (1979, 575–576)

Lewis thinks that it is "not good" to have two ways of classifying beliefs: by their content and the belief state the person is in. He claims that that would jeopardize the purpose of assigning contents to beliefs; of deciding *what* one believes. Having two components in the classification, as Perry does, is too complex, according to Lewis.

It might seem obvious that classifying anything using one parameter is simpler than using two. But it all depends, of course, on how we flesh out that one parameter. On Lewis's case, among other things, it involves accepting a very particular notion of "property," and the existence of possible worlds. According to most interpretations, it also involves a new and primitive attitude, self-ascription. Self-ascription, according to this interpretation, is primitive and it is a one-place relation: one self-ascribes a property. There is no two-place relation between the agent and the property. The property is in the head, in the agent's mind, and the agent simply self-ascribes it.

That Lewis is defending the existence of a special and primitive attitude, is a widely accepted claim. So much so, that it is often attributed to Perry as well, as if it were an integral part of any explanation of the role of indexicals and of indexically expressible beliefs. Magidor, for instance, considers that the following is a central tenet of “the myth of the *de se*,” defended by both Perry and Lewis:

There is a special class of propositional (or “propositional-like”) attitudes. These are *self-locating* or *de se* attitudes, ones that are typically expressed using indexical expressions such as “I” and “now” (call this claim “*Special Attitudes*”) (2015, 249)

This is simply wrong. It should be clear by now that Perry does not defend the existence of a special and primitive attitude. More surprising is his preferred interpretation of Lewis, according to which he didn’t either. According to this, on Lewis’s theory, “to say that one “self-ascribes” a property is simply a misleading way of saying that one is in a belief state the causal role of which is captured by the property. Self-ascription is not a newly discovered attitude” (2019, 132–33).

It is not easy to know whether this interpretation is accurate or not. Lewis’s undeniable originality and his idiosyncratic use of terminology make proper understanding difficult at times. It doesn’t help that his views have been discussed, interpreted and used very often and in various ways. But so have been Perry’s. I am inclined to give credit to Perry’s way of understanding Lewis’s notion of self-ascription, and Lewis’s theory in general. It does not contradict Lewis’s text, as far as I can tell, and it certainly makes Lewis’s theory more palatable.

The above paragraphs, of course, are not meant to be a presentation of Lewis’s proposal. My aim is just to argue that talk about a “Perry-Lewis” view on indexically expressible beliefs is misleading, and it very easily leads to confusions, attributing to one author the views of the other. For similar reasons, calling self-locating beliefs, and indexically expressible beliefs, in general, “*de se* beliefs” is not accurate, to say the least. Simplifying very much, *de se* beliefs is the term used by Lewis, who defined them as consisting in the self-attribution of properties. He claimed that all beliefs are *de se* beliefs. Perry does not believe that indexically expressible beliefs, and, in particular, self-locating beliefs, consist in self-attributions, or on any relation between the agent and a property (or any other entity, for that matter). He did not claim either that all beliefs are self-locating or indexically expressible beliefs.

Negationists like Cappelen and Dever, however, do not object to particular details of Perry’s and Lewis’s theories, but to what they both have in common: the need to modify the first tenet of the doctrine of propositions; that is, the claim that propositions —Fregean, singular, structured or possible-world— are not enough to classify beliefs; that beliefs do not *consist* in a relation with one proposition.

Cappelen and Dever are particularly worried about “perspectivism,” both in Perry and in Lewis. They think Lewis, Perry and all their “followers” cannot offer a notion of content “as an objective representation of the world” (2013, 173). But, clearly, Cappelen and Dever confuse self-ascription of properties (*de se*) or indexically expressible beliefs with the ascription of perspectival properties or the belief of essentially indexical contents. As Falk neatly puts it “It is a wonder, given their own appeal for objective properties in their last chapter, that they would not see the essential indexical is about locating oneself, not in a seeming world, but in the objective real world.” (2015, 428)

That Perry does not defend perspectivism in this sense is, I think, clear. In his book, Perry explains how Lewis’s proposal doesn’t lead to it either. He does so, mostly, on chapter 12, where he compares his views with Lewis’s, and offers a new and original interpretation of Lewis’s proposal. This is a key chapter of the book, and I believe Perry’s interpretation of Lewis deserves much more attention than what has received so far. If only because it makes Lewis’s proposal plausible and simpler than the usual interpretations. Also, because it makes it compatible with folk-psychology and the way we normally understand beliefs and our means to express them.

I’ve already said that Perry’s book answers to negationists’s criticisms. I hope I have also shown how this short book does much more than this. It does a great job at clarifying Perry’s claims in his classic 1977 and 1979 papers, answering many common misinterpretations; and it offers new and original proposals on, among others, Prior’s “thank goodness that’s over” paper and on Lewis’s theory. It does all this in a clear, well-informed and very entertaining way. This book should definitely be read by all interested in the philosophy of language and mind, and especially by those interested in issues related to the essential indexical and/or the *de se*.

References

- Boer, Steven E. & Lycan, William G (1980) “Who, Me?” *Philosophical Review* 89 (3):427–466.
- Cappelen, Herman and Josh Dever (2013) *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person.* Oxford: Oxford University Press.
- Castañeda, Hector-Neri (1966) “He’: A study in the logic of self-consciousness,” *Ratio* 8:130–157.
- Castañeda, Hector-Neri (1967) “Indicators and Quasi-Indicators,” *American Philosophical Quarterly* 4 (2):85–100.
- Castañeda, Hector-Neri (1968) “On the logic of attributions of self-knowledge to others,” *Journal of Philosophy* 65 (15):439–456.

- Chisholm, Roderick M. (1981) *The First Person: An Essay on Reference and Intentionality*. University of Minnesota Press.
- de Ponte, María (2017) "Promises, the present and now: Lessons from Austin, Prior and Kamp," *Journal of Pragmatics*, 112: 33–43.
- de Ponte, María and Kepa Korta (2017) "New thoughts about old facts. On Prior's root canal," in de Ponte & Korta (eds.) *Reference and Representation in Language and Thought*. Oxford: Oxford University Press, pp. 163–178.
- Devitt, Michael (2013) "The myth of the problematic *De Se*," in A. Capone & N. Feit (eds.) *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. Stanford: CSLI publications.
- Douven, Igor (2013) "The epistemology of *De Se* beliefs," in A. Capone & N. Feit (eds.) *Attitudes De Se: Linguistics, Epistemology, Metaphysics*. Stanford: CSLI publications.
- Falk, Arthur (2015) "Herman Cappelen and Josh Dever, *The Inessential Indexical: On the Philosophical Insignificance of Perspective and the First Person*". Oxford University Press, 2013. Xiii + 194 pages ISBN 978-0-19-968674-2," *Journal of the Indian Council of Philosophical Research* 32 (3):425–430.
- Falk, Arthur (Manuscript) "A defence of attitudes *De se* and *De nunc* from recent criticisms." Last version, 2015. Available at: <https://wmich.academia.edu/ArthurFalk>
- Holton, Richard (2015) "Primitive Self-Ascription: Lewis on the *De Se*." In Barry Loewer & Jonathan Schaffer (eds.), *The Blackwell Companion to David Lewis*. Blackwell.
- Lewis, David (1979) "Attitudes de dicto and de se." *Philosophical Review* 88 (4):513–543.
- Liao, Shen-yi (2012) "What Are Centered Worlds?" *Philosophical Quarterly* 62 (247):294–316.
- Magidor, Ofra (2015) "The Myth of the *De Se*," *Philosophical Perspectives* 29 (1):249–283.
- Millikan, Ruth Garret (1990) "The myth of the essential indexical," *Noûs* 24: 723–734.
- Ninan, Dilip (2016) "What is the Problem of *De Se* Attitudes?" In Stephan Torre & Manuel Garcia-Carpintero (eds.), *About Oneself: De Se Thought and Communication*. Oxford University Press.
- Perry, John (1977) "Frege on demonstratives," *Philosophical Review* 86 (4):474– 497. Reprinted in Perry (2000), pp. 1–27.

Perry, John (1979) "The problem of the essential indexical." *Noûs* 13 (1):3–21.
Reprinted Perry (2000), pp. 27–45.

Perry, John (2000). *The Problem of the Essential Indexical and Other Essays, Expanded Edition*. Stanford: CSLI.

Perry, John (2019) *Revisiting the Essential Indexical*. Stanford: CSLI Publications

Prior, Arthur N. (1959) "Thank Goodness That's over," *Philosophy* 34 (128):
12–17.

Prior, Arthur N (1968) "Now," *Noûs* 2 (2):101–119.

Stalnaker, Robert (2008) *Our Knowledge of the Internal World*. Oxford
University Press, Oxford.

CIP – Каталогизација у публикацији
Народна библиотека Србије, Београд

1

FILOZOFSKI godišnjak = Belgrade philosophical
annual / editor Voin Milevski. – God. 1, br. 1 (1988)– . –
Belgrade : Institute of Philosophy, Faculty of Philosophy,
1988– (Belgrade : Službeni glasnik). - 24 cm

Polugodišnje. – Glavni stvarni naslov od br. 28 (2015)
Belgrade philosophical annual. – Tekst na engl. jeziku.

ISSN 0353-3891 = Filozofski godišnjak

COBISS.SR-ID 15073792

BELGRADE PHILOSOPHICAL ANNUAL 35 (2)/2022
BELGRADE PHILOSOPHICAL ANNUAL 35(2)/2022

ARE THERE ESSENTIAL INDEXICALS?

Guest editor: Vojislav Božičković, University of Belgrade

John Perry	Are There Essential Indexicals?.....	7
Miguel Ángel Sebastián	Conscious Perception in favour of Essential Indexicality.....	13
Andrija Jurić	The Essential Indexical and Self-Consciousness: 'I', 'Now', and 'Here' as Aspects of Self-Consciousness	31
Peter Ludlow	Cognitive Dynamics: Red Queen Semantics versus the Story of <i>o</i>	53
Andreas Stokke	<i>De Se</i> Thinking and Modes of Presentation.....	69
Eduarda Calado Barbosa	For Their Eyes Only.....	89
Ernesto Perini-Santos		
María de Ponte	Demystifying the myth. Perry: Revisiting the Essential Indexical.....	107