



Semantic conventions and referential intentions

Jakub Rudnicki¹ 

Received: 14 June 2022 / Accepted: 9 June 2023 / Published online: 25 June 2023

© The Author(s) 2023

Abstract

According to intentionalism, the semantic reference of the uses of demonstratives is fixed, at least partly, by the speaker's referential intention. In this paper, I argue against the possibility of the existence of a semantic convention of this sort. My argument is placed in the Lewisian framework of signaling games and consists of several steps that correspond to four anti-intentionalist arguments, already present in the literature, that have proven inconclusive when employed separately and without being set in the mentioned framework.

Keywords Demonstratives · Intentionalism · Coordination · Context sensitivity · Reference · Semantics

1 Introduction

Intentionalism is the view according to which speakers' intentions play some role in the determination of semantic reference of uses or occurrences¹ of a given category of expressions. My main aim in this paper is to argue against this view. I will focus my discussion on the demonstrative "that" but I believe my argument naturally applies to any other semantic application of the intentionalist position as well, such as with regards to other indexicals, etc.

My argument is set in the Lewisian landscape of signaling games and the central lesson that I hope it gives rise to is that semantic convention of the form that intentionalism proposes has no chance of materializing and that, *a fortiori*, it is, therefore, an implausible analysis of our current semantics.

An additional minor aim of mine is to give new life to certain anti-intentionalist arguments that are already available in the literature but remain inconclusive insofar as

¹ I remain neutral between these two options, but in the rest of the paper I will assume uses to be the bearers of reference.

✉ Jakub Rudnicki
j.m.rudnicki@uw.edu.pl

¹ Faculty of Philosophy, University of Warsaw, Krakowskie Przedmieście 3, 00-927 Warsaw, Poland

they are not seen through the Lewisian signaling-game lens (and looked at separately rather than as elements of one, larger, complex argument).

In the next section, I will introduce the basic ideas behind what Borg (2006) called Convention Based Semantics, which is the Gricean–Lewisian project in the philosophy of language. In Sect. 3, I will briefly explain how semantic intentionalism sits within the Convention Based Semantics framework. In Sect. 4, I will describe the current state of the debate between intentionalism and anti-intentionalism² with regard to demonstratives and will provide the description of the aforementioned anti-intentionalist arguments. Finally, in Sect. 5, I will present my main argument aimed against the possibility of intentionalism.

2 Convention based semantics

Firstly, let me introduce the general ideas behind Convention Based Semantics. Within this approach, speakers' communicative intentions are a precondition for any form of meaning, and certain patterns in the execution of those intentions over time become conventional in the sense that they begin to be expected by the members of the community repeating those patterns to convey certain stable contents on behalf of their users. In other words, a supporter of Convention Based Semantics makes the natural claim that words and sentences can possess conventional meanings which are established by the way they are used with regularity by intentional creatures in order to convey their thoughts (Borg, 2006, p. 251).

The classic analysis of the first component of this duo was famously provided by Grice (1957) in his acclaimed definition of utterer's meaning. The Gricean tradition, as well as Grice himself at different times, is not in complete agreement as to the preferred exact formulation of the definition, but its crux, roughly, is that for the utterer to mean something (p , say) via a particular action, it is for them to intend to induce an effect (e.g. a belief that p) in their audience as a result of the audience's identification of this intention (which is therefore often referred to as the *reflexive* intention).

The most influential attempt to define the notion of convention, which constitutes the second part of the view of language offered by Convention Based Semantics, is thanks to Lewis (1969). We won't be very interested in the details of this definition here but I will present the general idea behind it because we will need to discuss one of the notions used in it. So according to Lewis, for a regularity in behavior in a particular type of situation to count as a convention, it needs to be common knowledge for the members of the community in which the regularity occurs that the community follows this regularity, and that its members expect each other to follow it, and that the regularity is a solution to a *coordination problem* arising in that type of situation.³

Coordination problems are situations in which in order to maximize outcomes/utilities, each of the agents involved needs to make choices that conform to those made by the others. So for example, if we want to meet at our favorite coffee

² By anti-intentionalism I shall mean any position according to which intentions play no role in the process of reference-fixing but that rather claims that this work is done by some external features of the context (or a combination thereof).

³ The complete definition, as originally formulated, can be found in Lewis (1969, p.76).

shop and have two possible days for that to happen, Thursday or Friday, we would both prefer a situation in which we come to the coffee shop on the same day, either both on Thursday or both on Friday. This creates the need for coordination. Otherwise, if we both come on different days, it creates an unwelcome outcome for all parties.

Formally speaking, coordination problems are games involving more than one Nash equilibrium⁴ between which the rational agents need to coordinate. In our scenario, the two equilibria are formed by pairs of strategies (*strategy profiles*) in which both of the agents decide to come on the same day as the other.⁵

Overall, in order to achieve this coordination, the agents will need to take into account the possible decision of the other agents involved. In our case, if I expect you to be at the coffee shop on Thursday/Friday, I would have to go there on Thursday/Friday as well.⁶ And likewise for you. In practice, coordination in a situation of this sort can be achieved in many different ways. We could both luckily turn out to come to the coffee shop on Friday for no explicit reason, simply hoping that the other will do the same; we might both know that we are very busy on Thursdays, and will therefore expect one another to be at the coffee shop on Friday; or maybe there is a general custom in our community to meet for coffee on Fridays that causes us to expect the other to visit the coffee shop on Friday; or most naturally, maybe we have simply directly agreed with each other to meet for coffee on Friday. In any case, if we manage to coordinate our meeting successfully, and if we were to find ourselves having a recurring need of coordinating coffee meetings like this, at some point, meeting on Fridays could become a mutually known convention that we could both refer to when deciding whether to show up at the coffee shop on Thursday or Friday.

For Lewis, communicative situations involving the use of signals are coordination problems in the same sense. So, for example, if it is agreed between us that you should leave the party (for example because of the confirmed presence of someone you'd prefer to avoid) if I secretly give you a sign with my right hand, and that you should stay (because of the confirmed absence of that person) if I use my left hand, then successful coordination takes place only if your action—leaving or staying—ends up matching my action—gesturing either with my right or my left hand, respectively—in the way just described. In such circumstances, the gesture of my hand counts as the signal.

But note that in principle, the connection between these two signals and the preferred actions of the hearer is purely arbitrary, and could just as well be reversed (i.e. the gesture with my right hand could prompt their staying, and with my left hand their leaving). Note also that in this case, proper full-blown coordination requires not only doing our part in the situation that actually happens at the party (the person that the hearer wants to avoid is either present or absent) but rather being able to do our part successfully irrespective of which of these possible scenarios materializes. These two observations suggest that the true objects of coordination in such signaling games are in fact fuller communicative strategies, i.e. modes of behavior sensitive, in the case of

⁴ I.e., informally, a combination of actions in which no agent would have been better off had he acted differently.

⁵ For simplicity, here and in later parts of the text, I am allowing only pure strategies.

⁶ This example assumes that none of the agents can come both on Thursday and on Friday but that they need to choose exactly one of those days.

the speaker, to all (in this case two) possible states of the world, and, in the case of the hearer, to all (here, we can assume again, two) signals. In other words, the actors need to coordinate how the speaker is going to code the state of the world that they observe with how the hearer is going to act upon being presented with the speaker's signal. In the remaining part of the paper, I shall assume that the relevant action on the part of the hearer shall be the forming of an appropriate interpretation of the speaker's signal, e.g. in a situation in which the signal is caused by the speaker's observation of state t , the appropriate action on the part of the hearer will be the forming of the interpretation that t holds. And, as far as we are interested here in referential expressions, if t is that object A is Q, then the relevant signal and interpretation will be one to the effect that it communicates something about A (rather than some other potential object B).

In a signaling game, just as in the coffee shop case, if we end up successfully coordinating the communicative strategies once, this fact can over time give rise to the stable convention according to which the gesture with the right hand is the signal for the hearer to leave the party, and the gesture with the left hand is the signal to stay (rather than vice versa).

To sum up, the general view that we have right now says that signaling or linguistic conventions for referring expressions stem from repeating needs to coordinate speakers' signals with hearers' interpretations. We have seen that such coordinating conventions, just as the coffee-shop-meeting convention above, may be born in different ways. Some will begin by accident, some will be grounded in cultural salience, some might result from an explicit agreement made in advance, etc. The common theme for them, though, is that in order for those one-time solutions to turn into fully-fledged conventions over time, the previous cases of success in coordination they resulted in must give rise to enough of a precedent to bring about the common expectation within the members of the community that those solutions will be replicated by them in the future. Once that happens for linguistic signals, the description of the community-wide convention effectively constitutes the description of the semantics of such signals within that community.⁷ Additionally, this ingrained evolutionary nature of the conventions suggests that we should expect them to change over time, and adjust to further, also potentially changing, needs of the population they serve and the changing circumstances in which they operate.

Combining both parts, the Gricean and the Lewisian, we can see that according to Convention Based Semantics, once the convention to refer to some object with the uses of a particular string of sounds, signs, or a gesture is set, the members of the population will naturally be choosing to convey the referential element of their communicative reflexive intention, via which they want to induce in the hearer the belief that they want to refer to the particular object,⁸ through the uses of the very sounds, signs or gesture that are conventionally used to do it.

⁷ In other words, on this approach, the semantics of the signals used within a community can only be described at the community-wide level as it stems from the existence of the general conventions governing the uses of the signals. And conventions, by definition, transcend individual speakers or pairs involved in one-off exchanges. I would like to thank an anonymous referee for recognizing the need for this clarification.

⁸ The "particular object" can also be, of course, the object playing a particular role in the context, etc.

3 Intentionalism within the convention based semantics

The most typical approach towards referential expressions within the Convention Based Semantics framework has it that the conventional semantic characteristics of word-types (like the Kaplanian (1989a) characters) take them all the way to their semantic reference or, for cases of indexical expressions, they get a little assistance from (preferably) easily accessible features of contexts such as who the speaker is, the time or place of the utterance, etc.

For demonstratives in particular, there have been different proposals offered as to what this relevant, reference-determining feature of context could be: demonstrative gestures (McGinn, 1981; Rudnicki, 2022, 2023), being at the center of attention in a discourse (Stojnić et al., 2017), and potentially others of this sort,⁹ but the one I am going to focus on, given the topic of this paper, are speaker's intentions. Schematically speaking, on the intentionalist view, the reference of a token of "that" is fully determined in the following, straightforward way:

character of "that" + speaker's intention (as a semantically relevant feature of context)->refere and the classic proponent of such a view is Kaplan (1989b):

In *Demonstratives* I took the demonstration (...) to be criterial for determining the referent of a demonstrative.

(...)

I am now inclined to regard the directing intention, at least in the case of perceptual demonstratives, as criterial [...].

On my current view, the referent of a true demonstrative is determined by the utterer's intention. (582, 585)

More recently, versions of this view were proposed also by Åkerman (2009), Radulescu (2019), Stokke (2010), and Viebahn (2020).

One clarification that needs to be made is that the purely semantic version of intentionalism above should be distinguished from the other, more pragmatic, version of a view that also appeals to speakers' intentions when it comes to demonstrative reference. This other view, though, supported by Bach (2017), Heck (2014), and Smit (2012), contrary to its above semantic cousin, claims that the uses of demonstratives do not refer as a matter of the semantics of these expressions at all, but rather that it is only the users of such expressions that do, and that this phenomenon is therefore pragmatic in kind. This also means that from this perspective, due to semantic underdeterminacy, the resolution of reference will often involve abductive, open-ended considerations, rather than purely algorithmic ones stemming from the rules provided by the semantics of the relevant expressions (as the above, semantic version of intentionalism predicts). A succinct description of the view is provided by Bach (2017):

I will argue that there is no good reason for maintaining that demonstratives refer in their own right. They do not have meanings that determine their reference as a function of context. And the sort of referential intention that can plausibly be

⁹ Also on the table are anti-intentionalist proposals by Wettstein (1984) and Gauker (2008).

attributed to an ordinary speaker is not semantically [but only pragmatically – AU] relevant.

(...)

The natural alternative to the claim that demonstratives have meanings that determine their reference as a function of context is that their meanings merely constrain their literal use. Of course, speakers use them to refer, but this can be explained without attributing references to demonstratives themselves. (58)

I won't have too much to say about this view in this paper, and my argument developed below should not be taken as applying to it. Having said that, it will make a few brief reappearances in the latter parts of the paper. (A terminological sidenote: I shall use *intentionalism* throughout the paper to refer to the main, semantic position that interests me in this paper. I shall also sometimes call it *semantic intentionalism* to highlight the contrast with the currently discussed view—that I shall call *pragmatic*—when called for.)

4 The current state of debate on demonstratives

The general debate between intentionalists with regards to demonstratives and their opponents has been centered around a couple of arguments and attacks of the latter group on the position proposed by the former, and their reactions to them. The main one, as is typically the case in semantic debates, revolves around the referential intuitions raised by particular scenarios. Here, the ones that have gathered the most attention are situations in which the speaker's act of pointing (or other contextual cues) suggests a different interpretation of their uttered demonstrative than their intention does, or in other words, there happens to be a mismatch between the object of the speaker's intention and the object pointed at. One example of such a mismatch was proposed by Gauker (2019):

Imagine that we are seated at a table in a restaurant. I look at the fork next to my plate. I point at it with my finger. I utter, "That is dirty". However, as I speak, I am thinking of the fork in front of the person sitting to my left. A moment earlier I had looked at it and seen that it was dirty. My glance to my own fork and my gesture with my finger were entirely accidental. I did not mean to draw attention to the fork in front of me. What I intended to speak of—intended to refer to—was the dirty fork in front of the person to my left. The fork in front of me is perfectly clean. (109)

Such scenarios are usually taken to be raising non-intentionalist intuitions, as it seems that, intuitively, the speaker has in fact referred to the object he mistakenly pointed at rather than the one he intended to refer to. Below, I shall call this sort of argument the *cases of mismatch argument*.

Furthermore, given that in principle we have two types of possible situations, ones involving the mismatch, and the other, typical ones in which the mismatch does not take place, there is a worry that speakers' intentions are essentially redundant when it comes to the fixing of reference since in typical cases they simply predict the same

results as the contextual cues, and in the non-typical mismatch cases, the intuitions are in line with the predictions stemming from the cues and not from the intentions (let me call this part of the argument the *anti-intentionalist redundancy* argument).

Even though the above argument combining the cases of mismatch with the redundancy suggestion may seem compelling at first, the intentionalists have a pretty straightforward reply at hand. Namely, they can simply argue that the relevant intuitions raised by the scenarios in question track not the objective reference of the use of the demonstrative but rather what the hearer could reasonably take to be that reference (Åkerman, 2009; Radulescu, 2019).

The next argument, made by Gauker (2008), which I shall call the *unavailability argument*, makes the point that given the fact that the speakers' intentions are not directly available to the hearers, a semantics based around them would make it seem mysterious that demonstrative communication is typically successful. The reason for this is that there would seem to be no reliable way for the hearers to interpret the speakers' intentions based on what is available to them (and independently of the knowledge of the propositions expressed by the speakers).

But here, again, the intentionalists can counter that this point confuses the epistemic claim about what is available to the hearers and could be reliably used by them for the interpretation of the utterance, with the metaphysical claim about what, in fact, determines reference.

The last anti-intentionalist argument that I want to discuss, similarly to the extension of the cases-of-mismatch argument, claims a redundancy with regards to the semantic, reference-fixing potential of speakers' intentions (I shall call it the *anti-reference redundancy* argument). It does so in a different way, though. Here, the observation is that, just as in the widely accepted Gricean model introduced above, the interpretation of the communicative intentions of the speaker is the crucial element of any communication, and not only in cases involving the use of demonstratives. Therefore, the argument follows, the imposition on the hearers of the requirement to interpret the reference-fixing intention of the speakers is superfluous given that they need to interpret the general communicative intention anyway (Heck, 2014; Smit, 2012).¹⁰ This would suggest that the rightful place for speakers' intentions is outside of the semantically relevant context.

But this argument has also been reacted to. King (2013, p. 166) noticed that it rests on a certain conception of the semantics-pragmatics interface according to which the reference of demonstratives ends up being resolved by the pragmatic side of things. Once this conception is rejected, one can straightforwardly double down on claiming that the speakers' relevant intentions are part of the semantically relevant context that plays a part in the process of reference fixing.

As we can see, the debate between the competing positions seems to be locked in a deadlock. Even though I am generally sympathetic to the above anti-intentionalist arguments, I have to admit that the intentionalists' reactions to them are difficult to reject without begging the question against their position. In the next section, I hope to

¹⁰ Smit (2012, pp. 54–59) provides also a second version of similar considerations suggesting that intentionalists effectively commit a category mistake stemming from the fact that the referential convention they posit does not rationalize any action that wouldn't already belong to the basic principles of communication and optimal epistemic conduct, and that it therefore does not constitute a convention at all.

break this tie by placing the anti-intentionalist arguments within the Lewisian signaling games framework in which the ad hoc nature of the above reactions to the arguments becomes clear.

In the context of the debate we are interested in, the main benefit that the Lewisian framework brings to the table is that it imposes a constraint on what can count as semantic —namely that it must be conventional.¹¹ This is relevant because of the nature of the intentionalist replies which boil down to suggesting the merely pragmatic relevance of the anti-intentionalist arguments, and placing what is supposed to be semantically relevant outside of the reach of those arguments. The constraint, together with the evolutionary approach to how conventions are born, will prove helpful in showing how the two anti-intentionalist reactions to the unavailability argument possible within the Lewisian framework fall prey either to the anti-intentionalist redundancy argument (proving that for the semantic convention resulting from one of the reactions the real reference-fixing role is performed by the contextual cues rather than by the referential intention) or the anti-reference redundancy argument (yielding the version of intentionalism resulting from the second reaction pragmatic).

5 The argument against the possibility of intentionalism

Before moving forward, let us briefly take stock. We started with introducing the basic ideas behind what I've called, following Borg, Convention Based Semantics. It is an approach to the nature of meaning that rests on two pillars: Grice's understanding of utterer's meaning and Lewis's (or, more precisely, inspired by Lewis) understanding of conventional meaning as a common-knowledge strategy of solving the recurring need to coordinate. I have also tried to emphasize the evolutionary, changing nature of the latter, and its inception from the former.

I have also characterized the intentionalist approach to the semantics of demonstratives based on the idea that their uses determinately refer on the basis of the Kaplanian characters of their types operating on speakers' intentions playing the role of the relevant feature of context.¹²

In a second I shall move to the argument for why I think intentionalism is not a feasible semantics for demonstratives, motivating this claim with the observation that the semantic convention governing the uses of these expressions could never evolve into what its proponents suggest it looks like. The last thing that I would like to underline before moving forward is that the success of the argument that follows

¹¹ An anonymous referee has made the remark that given that many authors think of semantics as conventional without being motivated by considerations having to do with signaling games, I might be employing too heavy a machinery to establish an otherwise commonplace result. The referee has also mentioned Smit's (2012) and Devitt's (2021) work as highlighting the underappreciated role of conventionality as a constraint on semantics.

¹² I use the notion of the 'relevant' feature of context rather than the 'Kaplanian' or 'formal' context because there are at least two ways of formally expressing the discussed semantic idea in the Kaplanian framework. One could either treat the intention as the feature of the Kaplanian context, as Stokke does, or one could leave this role directly to the *demonstratum* (should we call it *intentionatum* not to raise the unwanted connotation with demonstrative gestures?) and say that it is simply picked out by the intention lurking in the wider, non-formal, context (cf. Kaplan, 1989a).

is not dependent upon the correctness of the Gricean-Lewisian picture—described above—in its entirety. Specifically, it does not depend upon the details of the exact analysis of the notions of convention or utterer’s meaning as well as any particular claim about whether and how utterer’s meaning gives rise to semantic conventions. What my argument does need is merely the widely plausible assumption that semantics is tied to conventionality and the aforementioned evolutionary approach towards it that allows for assessing communicative strategies based on whether they could bring about the common expectation within the members of the community that those coordination solutions would be followed by them in the future.

Now, let’s begin the presentation of the argument by looking more closely at a simple signaling game that includes pointing as the referential signal.¹³

Assume the following: we have two agents, S(peaker) and H(earer); two objects of potential reference, A and B; two potential states of the world, $t_1 = A$ is Q, $t_2 = B$ is Q; two potential referential signals, s_1 and s_2 (which are supposed to differentiate only between A and B; the rest of the proposition, “is Q”, is outside of our interest); and it is common knowledge that S always knows whether t_1 or t_2 and that she wants to honestly inform H whether t_1 or t_2 (i.e. given the remark in the previous brackets, rather that she wants to talk about A or B). Finally, let’s also assume the following symmetric utilities for both agents: $U(t_1, i_1) = 1$; $U(t_2, i_2) = 1$; $U(t_1, i_2) = 0$; $U(t_2, i_1) = 0$; where i_1 and i_2 are the two possible interpretations on the part of the hearer (either that A is Q or that B is Q, respectively).

A brief description of the formal nature of this game is as follows. Each agent has four potential (pure) strategies. For S, those coding strategies (functions $T \rightarrow SI$, from the set of states of the world T to the set of signals SI) come from the different ways in which she can tie the signals to the states of the world ($S_1: \{ \langle t_1, s_1 \rangle ; \langle t_2, s_2 \rangle \}$, $S_2: \{ \langle t_1, s_1 \rangle ; \langle t_2, s_1 \rangle \}$, $S_3: \{ \langle t_1, s_2 \rangle ; \langle t_2, s_2 \rangle \}$, and $S_4: \{ \langle t_1, s_2 \rangle ; \langle t_2, s_1 \rangle \}$). And for H, the interpretative strategies (functions $SI \rightarrow I$, from the set of signals SI to the set of interpretations I), which are the different ways in which he can connect his two available interpretations with the two potential signals that he can get from S ($H_1: \{ \langle s_1, i_1 \rangle ; \langle s_2, i_2 \rangle \}$, $H_2: \{ \langle s_1, i_1 \rangle ; \langle s_2, i_1 \rangle \}$, $H_3: \{ \langle s_1, i_2 \rangle ; \langle s_2, i_2 \rangle \}$, and $H_4: \{ \langle s_1, i_2 \rangle ; \langle s_2, i_1 \rangle \}$). This gives us sixteen potential strategy profiles, i.e. S-strategy-H-strategy pairs. Six out of them are Nash equilibria but only two are *separating* Nash equilibria (i.e. Nash equilibria in which there is a 1–1 correspondence between the states of the world and the signals¹⁴). And these are, unsurprisingly, $P_1: \langle S_1; H_1 \rangle$ and $P_2: \langle S_4; H_4 \rangle$. In what follows, we shall therefore focus only on these two.

Given the above assumption that the signal is supposed to be a pointing gesture, let’s think of s_1 as pointing at A, and of s_2 as pointing at B. Therefore, the two available signaling profiles that the agents need to coordinate between are such that on P_1 , pointing at A refers to A, and pointing at B refers to B; and on P_2 , pointing at A refers to B, and pointing at B refers to A. If one of these profiles dominates the other, given

¹³ For a discussion on the topic of signaling games, see e.g. van Rooy (2004).

¹⁴ The additional interesting quality of separating Nash equilibria in these kinds of games with the same number of states, signals, and interpretations is that they and only they will be evolutionarily stable (Wärnerd, 1993). This happens to be potentially an important property justifying our focus on only these two profiles and ignoring the other four equilibria, but I won’t be getting into this matter here.

the power of precedent, etc., it has the potential to turn into a full-blown referential signaling convention stably used in this population in cases of this sort. And again, just as we have seen before, there are several ways in which one of the profiles could get a head start thanks to the first successful exchange(s) achieved when following it. The successful coordination might be due to a lucky decision on the part of the agents; the salience of the object the speaker ‘referred’ to might have suggested to the hearer that she is following, say P_1 (if the salient object was pointed at), rather than P_2 (in which case she would point at the non-salient object), and so on. What is very important though, is that for that action of pointing to have the chance of obtaining the referential conventional meaning of referring to the object pointed at (or to the object other than the one pointed at, as in the other profile), it needs to be possible for all the agents involved to figure out what that relevant signaling action is. And in the case of pointing this is (at least in principle) easily met. For example, the hearer realizes that the speaker is trying to refer to a particular object (say, either luckily or because of the object’s salience), he sees the action of pointing, and once his interpretation of the signal is later confirmed, he can expect that signal to be used in the same way in the future.

And here is where we get to the first problem of intentionalism. We just saw how a convention could arise on which pointing gestures are the signaling actions that fix reference. Note now that on intentionalism, i.e. the position according to which it is the speaker’s intention of the relevant sort that fixes semantic reference, the referential signaling action is the action of forming this relevant intention (i.e., say, the action of forming the intention for the token of “that” to refer to A in t_1 , and the action of forming the intention for the token of “that” to refer to B in t_2). But this action is not directly available to the hearer! And this is why we should conclude that any convention built around such a signaling action could not bootstrap itself in the same natural way that we saw the pointing convention doing. In other words, even if the hearer ended up interpreting the speaker’s ‘utterance’ successfully, given that the speaker’s action of forming the relevant intention is not available to the hearer, the hearer will not be able to expect such an action to fix the reference in the future cases of communication, simply because they have not registered it as a signal in the original situation.

But this observation, which by the way constitutes the analogue of Gauker’s unavailability argument from before, need not be conclusive. That is because one could make the perhaps plausible point that even though there is the clear issue of the impossibility of the spontaneous bootstrapping of the intentionalist convention, it still does not eliminate the possibility of the emergence of some sort of reliable connection between some contextually available cues (one could think of pointing as a possible one, but the idea here is that in principle this could be *anything* available to be read off by the hearers, including several different cues at once) and the speakers’ intentions that the speakers could utilize and which could guide the hearers’ interpretations.

In principle, there are two types of such possible connections: loose and strict. A loose connection would be one offering no direct determinate way for the hearer of arriving at the speaker’s reference based only on this set of cues. Hearers would have to at least partially resort to wider considerations regarding the possible motivations of the speakers, etc.

A strict connection, on the other hand, would offer a determinate way of figuring out the speakers' intentions based solely on the contextually available cues. This would essentially amount to a definite algorithm tying the cues to the intentions.

But out of these two solutions to the problem of the possible emergence of the intentionalist convention, the first one, centered around the loose connection, is not available to the defender of this view. The reason for this is that the appeal to the loose connection between the cues and the intentions that needs to be supplemented with the mentioned wider considerations is clearly out of reach of the proponent of the semantic version of intentionalism that we are focused on here. That is because potentially open-ended considerations regarding the possible motivations of the speaker and so on are definitely pragmatic in nature, and therefore break the requirement characteristic of the semantic version of intentionalism to the effect that the conventional signal ties with the reference without any pragmatic mediation. Taking such a step would result in the collapse of the semantic version of intentionalism into its pragmatic cousin, and is therefore of no help to the view under consideration. This observation, by the way, constitutes the analogue of the anti-reference redundancy argument from before.¹⁵

Before we consider the strict connection possibility, this is the right moment to look at how the above discussion applies to the real-world exchange between Gauker (2008) and Åkerman (2009).¹⁶ As I have alluded to in Sect. 4, Gauker's problem with intentionalism is that speakers' intentions are not directly available to the hearers. In my restatement of this observation, it points to a problem with the possibility of the bootstrapping of an intentionalist convention due to the fact that it is not clear how the action of forming the intention could be expected by the hearer to play the role of the signal in the future.

This is when Åkerman's reply, also mentioned in Sect. 4, comes into the picture. As we remember, the reply was that the unavailability observation incorrectly identifies what could be reliably (but, according to Åkerman, non-algorithmically) used by the hearers for the interpretation of the utterance, with what, in fact, determines reference (2009, p. 158).

From the perspective of this paper, there are two ways in which this reaction can be developed. The first, which is the one that coincides with Åkerman's actual view, is to treat the non-algorithmic cues as completely separate from the content of the semantic convention governing the demonstrative reference. The second would be to try to find a place for those cues as conventionally connected to speakers' intentions.

¹⁵ An anonymous referee asks what the consequences of this point could be for Gauker's (2008) anti-intentionalist view as it might be read as non-algorithmic but nevertheless semantic (rather than pragmatic). First, let me clarify that the way I'm using the labels of semantic and pragmatic intentionalism (or in the case of Gauker—anti-intentionalism) is supposed to mirror Borg's (2006) distinction for A-style and B-style Convention Based Semantics. According to Borg, this distinction comes down to whether a particular approach in its description of the retrieval of content, alongside involving an algorithmic element, appeals also to an open-ended, inferential component (Sect. 12.3). I think that my usage of the notions of "semantic" and "pragmatic" in this paper to mimic this distinction is justified by the fact that the views by Bach, Heck, and Smit that fall naturally into Borg's B-style Convention Based Semantics explicitly appeal to the pragmatic notion of reference. Given the above, it is by definition the case that any appeal to a non-algorithmic mechanism will yield a view "pragmatic" according to how I am using this notion here. Having said that, it does not mean that yielding a view pragmatic makes it inferior in any way. As I explained on p. 6, my argument should not be understood as applying to pragmatic views at all.

¹⁶ I would like to thank an anonymous referee for pressing me to clarify the connection here.

But we have already seen what is wrong with both of these approaches. The former capitulates right away when faced with the unavailability observation as it does not even attempt to look for a conventional tie between the cues and the speaker's intention. The latter, on the other hand, given the non-algorithmic nature of the cues discussed by Åkerman,¹⁷ repeats the fate of the loose connection solution and yields the proposal along these lines pragmatic in nature.

Let's now move to the discussion of the strict connection possibility. The first thing to notice about it is that it brings about a new coordination problem that needs to be solved, and therefore, in the end, a need for a new subsequent convention. This coordination problem is of course one about how the agents should be tying the two different signals, i.e. speakers' intentions, with the contextually available cues. For the speaker, since she is trying to convey her relevant intention, this problem comes down to the choice of the function from signals to cues, while for the hearer, this comes down to the choice of the function from the cues available to the semantically relevant signals, i.e. speaker's intentions. So for example, one of such strategies for the speaker could be (for c_1 and c_2 being the only two available arrangements of cues in the set of cues C): $\langle s_1; c_1 \rangle$, $\langle s_2; c_2 \rangle$; and for the hearer: $\langle c_1; s_1 \rangle$, $\langle c_2; s_2 \rangle$. Here, successful coordination will be achieved if and only if the resulting strategy profile is composed of two functions that are inverses of each other (below, I shall work with the assumption that the above pair of strategies becomes the dominating profile).

Overall then, rather than, as we had previously, the single functions $T \rightarrow SI$ for the speaker, and $SI \rightarrow I$ for the hearer, we now have a general picture in which the full strategies for the speaker and hearer are respectively $T \rightarrow SI \rightarrow C$, and $C \rightarrow SI \rightarrow I$.

But this raises the question of what exactly are the functions $SI \rightarrow C$ and $SI \rightarrow I$ doing in this picture, given that we could represent both of these strategy schemas as compositions of the two functions involved in each, i.e. finally as only involving single functions: $T \rightarrow C$ for the speaker, and $C \rightarrow I$ for the hearer?

And this is not only formally so, but philosophically, too. If we watch what the agents under such a "double-function" convention are doing, we will see speakers trying to provide cues that match the state of the world that they are trying to report, and hearers trying to produce interpretations matching the cues they are provided with. This suggests that the assumed signals, i.e. speakers' intentions, are not playing any part in the communicative process allegedly governed by the convention that treats them as relevant semantic signals. This observation, by the way, constitutes the analogue of the anti-intentionalist redundancy argument from before.

Furthermore, note that this observation holds even for cases in which the speaker mistakenly acts according to a different (other than the dominating one that she is trying to conform to) signaling system when she performs the cues that are supposed to guide the interpretation of her intention, i.e. e.g. the cues actually performed are commonly taken to signify the other potential signal/intention than the one present in this particular situation.

In such a situation, it is still the case that the speaker was trying to match the cues with the state of the world that she observed, and that the hearer will try to match his

¹⁷ Åkerman, or any other intentionalist, could of course, in principle, opt for a strict connection solution too. I return to the discussion of this possibility in the next paragraph.

interpretation with the cues provided. This clearly suggests the primacy of the cues over the alleged intentions-signals. Also, this case shows that, and has a simple explanation as to why, if the hearer is led astray as a result of the speaker's mistake, it is clearly the speaker who is at fault because she has not conformed to the right strategy tying the cues with the states of the world but rather acted in accordance with another. This is also true if we retain the double-function schema since it is still the case that the speaker will have to be judged as wrongly applying the SI→C part of her complete strategy. And this, again, proves the primacy of the cues over the intentions. This reasoning, by the way, constitutes the analogue of the original cases of mismatch argument.

Also, note that the answer that the hearer could be able to see through the speaker's mistake and end up interpreting him successfully, and therefore that it is the intended object that is objectively referred to after all, is, again, not available to the proponent of semantic intentionalism, as it appeals to the loose, pragmatic, connection between the cues and the intentions and therefore is still not a case in which the interpretation is driven by the algorithmic intentionalist semantic convention.

(The above line of reasoning is also why I think that my overall argument is resistant to the observation that the shape of semantics could be more directly affected by biological (rather than rationalistic) factors (such as e.g. universal grammar) and therefore that my conclusions are at best correct only conditionally on that not being the case. The above argument is general—it shows that, no matter what the provenance of the semantics (at least according to how I am using this notion here, i.e. as a feature of a conventional signal), it is simply practically impossible for the agents to be following an intentionalist convention at least as long as they do not possess direct mind-reading abilities.)¹⁸

[Finally, I also do not think that for my argument to reach the intended anti-intentionalist conclusion, it needs to assume a particular, Lewisian, story about the nature of coordination. An anonymous referee suggests that thinking about coordination in terms of coordinating mental representations or mental states, along the lines suggested by Cumming (2013a, 2013b), could perhaps turn out friendlier from the perspective of the intentionalist. Even though I agree with the referee that one could certainly think about coordination differently than I (following Lewis) do and that Cumming's proposal should be taken as a possibility, I am hesitant to acknowledge that the choice I have made has the tendentious consequences the referee implies. The referee's suggestion is that given Cumming's understanding of coordination as coordination of mental states, one could think of those mental states (at least on the part of the speaker) as (something like) referential intentions, and that this would almost by definition make them not redundant in the way I am trying to argue because the meaning of the public signal used to convey the content of these coordinated mental states could only be given by reference to these very states.

I think I can grant all of that. In fact, as a supporter of Convention Based Semantics as described above, I do believe that mental states are central to meaning in general. But note that the main question that interests me in this paper is not *what needs to be appealed to when trying to tell the meaning of a particular signal*, but rather *what is*

¹⁸ I would like to thank an anonymous referee for pressing me on this issue.

the relevant signal in the first place (i.e. whether the referential intention can be a part of the signal alongside the vocalized demonstrative).

In other words, my claim is not that all mental states are communicatively redundant in the sense that they never shed light on any features of communication but rather only that they are redundant (and generally not fit for the job for the other reasons that I discuss) as far as they would be supposed to play the role of signals. And note that the Cumming story supported in the referee's remarks does nothing to establish the role of signals for the mental states it takes to play the central role in the process of coordination. This is explicitly supported in the following passage from Cumming (2013a):

The speaker's strategy of referential coordination is for his [mental - AU] symbol a to refer to some object x and simultaneously to corefer with its public expression σ . The hearer's strategy is for her [mental - AU] symbol b to refer to y and to corefer with the expression it is used to construe—also σ . (p. 394)

Finally then, I think that even if I assumed the picture of coordination offered by Cumming, the main result, i.e. that the referential intentions could not play the role of signals, would remain intact.]

To sum up, my whole argument has the following structure. First, in line with the unavailability observation from Sect. 4, we noticed that the fact that speakers' intentions are not directly available to the hearers poses problems for any semantic convention centered around such intentions to bootstrap.

This was shown not to be conclusive against the possibility of the emergence of an intentionalist convention as it is still possible that there appears some connection between speakers' intentions and some cues contextually available to the hearers that allow them to reliably read off the former. Such connections can belong to one of two general types: loose or strict.

The loose connection, i.e. the one that is not algorithmic in nature but involves wider, open-ended reasoning on the part of the hearer, is not available to the proponent of semantic intentionalism, as it clearly allows the pragmatic intrusion into the process of reference-fixing. Therefore, on this attempt, intentionalism simply collapses into its pragmatic version. This point is in line with the anti-reference redundancy argument from Sect. 4.

Finally, we observed that the strict connection hypothesis cannot help the semantic intentionalist theorist either. The reason was that this solution requires an additional coordination strategy tying the alleged intentional signals to the available cues, and once they are set, the signals become completely redundant as all of the relevant semantic and communicative work is done by the cues. This observation is the analogue of the anti-intentionalist redundancy argument from Sect. 4.

Furthermore, this claim about the speakers trying to match the cues with the state of the world and the hearers trying to make their interpretations follow the cues provided, happens to be true even in cases in which the speaker mistakenly ends up following the wrong strategy. In other words, even when the speakers misalign the cues with regard to the state of the world, it is still the will to perform the right cues that drives the speaker's signaling behavior and the will to interpret them correctly that drives the hearer's interpretative behavior. And given that the shape of the ruling convention

is essentially defined by what people following it do, it has to be concluded, in line with the cases of mismatch argument, that the convention which is in fact followed on the strict-connection extension of intentionalism is an anti-intentionalist one, and therefore, that the reference is the object selected by the external contextual cues rather than by the alleged intention-signals.

6 Conclusions

In this paper, I argued for the implausibility of the intentionalist semantic position with regard to demonstrative expressions. According to this view, these expressions refer purely as a matter of their semantics and of the semantically relevant feature of the context—the speaker’s intention.

I have outlined the main anti-intentionalist arguments present in the literature and showed why they are not conclusive. My own argument rests on the idea of transporting the debate to the realm of Lewis’s signaling games, and on taking a closer look at what it takes for a particular referential strategy to be able to become a semantic convention in a particular population.

I believe that my argument shows conclusively that the convention promised by intentionalism simply cannot arise as one practically followed by the members of any population who lack direct mind-reading ability. What I take to be an additional elegant result of my argument is that it is essentially a conglomerate of the particular mentioned arguments taken from the literature: the unavailability argument, the argument from the cases of mismatch, and the two varieties of the redundancy argument. By putting those arguments together, and setting them in the mentioned Lewisian landscape, I was able to give them new life and construct out of them a conclusive, coherent refutation of intentionalism.

Acknowledgements My work on this article was funded by a National Science Centre, Poland, Grant under award 2019/35/N/HS1/02009. I would like to express gratitude to two anonymous referees for this journal whose comments have helped me significantly improve the paper. I would also like to thank Tadeusz Ciecierski, Dominik Dziedzic, Paweł Grabarczyk, Piotr Makowski, Maria Matuszkiewicz, Joanna Odrowąż-Sypniewska, and Piotr Wilkin for discussions on this material. My special thanks go to Genoveva Martí, Alex Radulescu, and Indrek Reiland for sharing with me their thoughts on an early draft of this paper.

Funding This study was funded by a National Science Center, Poland, Grant under Award Number 2019/35/N/HS1/02009.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use

is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Åkerman, J. (2009). A plea for pragmatics. *Synthese*, 170(1), 155–167.
- Bach, K. (2017). Reference, intention, and context. In M. de Ponte & K. Korta (Eds.), *Reference and representation in thought and language* (pp. 57–72). Oxford University Press.
- Borg, E. (2006). Intention-based semantics. In E. Lepore & B. Smith (Eds.), *The Oxford handbook to the philosophy of language* (pp. 250–266). Oxford University Press.
- Cumming, S. (2013a). Creatures of darkness. *Analytic Philosophy*, 54(4), 379–400.
- Cumming, S. (2013b). From coordination to content. *Philosophers' Imprint*, 13(4), 1–16.
- Devitt, M. (2021). *Overlooking conventions: the trouble with linguistic pragmatism*. Springer.
- Gauker, C. (2008). Zero tolerance for pragmatics. *Synthese*, 165(3), 359–371.
- Gauker, C. (2019). Against the speaker-intention theory of demonstratives. *Linguistics and Philosophy*, 42(2), 109–129.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 66(3), 377–388.
- Heck, R. (2014). Semantics and Context-dependence: Towards a Strawsonian account. In B. Sherman & A. Burgess (Eds.), *Metasemantics: new essays on the foundations of meaning* (pp. 327–364). Oxford University Press.
- Kaplan, D. (1989a). Demonstratives: An essay on the semantics, logic, metaphysics and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes From kaplan* (pp. 481–563). Oxford University Press.
- Kaplan, D. (1989b). Afterthoughts. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 565–614). Oxford University Press.
- King, J. C. (2013). Kent bach on speaker intentions and context. *Croatian Journal of Philosophy*, 13(2), 161–168.
- Lewis, D. K. (1969). *Convention: A philosophical study*. Wiley-Blackwell.
- McGinn, C. (1981). The mechanism of reference. *Synthese*, 49(2), 157–186.
- Radulescu, A. (2019). A defence of intentionalism about demonstratives. *Australasian Journal of Philosophy*, 97(4), 775–791.
- Van Rooy, R. (2004). Evolution of conventional meaning and conversational principles. *Synthese*, 139(2), 331–366.
- Rudnicki, J. (2022). Speaker's intentions, ambiguous demonstrations, and relativist semantics for demonstratives. *Philosophia*, 50(4), 2085–2111.
- Rudnicki, J. (2023). Can the reference of a use of "That" change? Assessing nonstandard approaches to the semantics of demonstratives. *Journal of Pragmatics*, 209, 31–40.
- Smit, J. P. (2012). Why bare demonstratives need not semantically refer. *Canadian Journal of Philosophy*, 42(1), 43–66.
- Stojnić, U., Stone, M., & Lepore, E. (2017). Discourse and logical form: Pronouns, attention and coherence. *Linguistics and Philosophy*, 40(5), 519–547.
- Stokke, A. (2010). Intention-sensitive semantics. *Synthese*, 175(3), 383–404.
- Viebahn, E. (2020). Ways of using words: On semantic intentions. *Philosophy and Phenomenological Research*, 100(1), 93–117.
- Wärneryd, K. (1993). Cheap talk, coordination, and evolutionary stability. *Games and Economic Behavior*, 5(4), 532–546.
- Wettstein, H. K. (1984). How to bridge the gap between meaning and reference. *Synthese*, 58(1), 63–84.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.