Sellars and Davidson on meaning-matching

Christopher Gauker

Department of Philosophy, Faculty of Social Sciences,

University of Salzburg, Salzburg, Austria

This is the prepublication version. The paper was published in Willem A. de Vries and Marc A. Joseph, eds., *Sellars and Davidson in Dialogue*, Routledge, 2025, pp. 33–55.

Abstract:

Both Sellars and Davidson sought recipes for explaining the meanings of linguistic expressions. For Davidson, the meaning of a sentence is given by a statement of its truth conditions derivable from a general theory of truth for the language. However, when Davidson undertakes to explain each speaker's understanding of the sentences of his or her own language, the theory seems to dissolve into a formless theory of interpretation. Davidson never confronts the nature of basic linguistic communication, because he mistakenly assumes that languages are not essentially shared. For Sellars, the meaning of an expression is explicable in terms of the rule-governed moves and transitions that it is involved in. Sellars does not clearly explain the nature of these moves and transitions, and he does not provide any basis for evaluating them. He regards thoughts as modeled on speech, but only on nondialogic self-directed speech. Consequently, he too neglects to explicate the nature of basic interpersonal linguistic communication. In closing it will be contended that we do not need a theory of meaning in order to have a theory of basic linguistic communication. Basic linguistic communication can rest on a foundation of imagistic cognition, and talk of meaning serves a regulatory role that guides the conduct of linguistic communication.

Introduction

So far as I am aware, I am the only person living or dead who was ever enrolled in classes of both Donald Davidson and Wilfrid Sellars. Davidson directed my bachelor's thesis at the University of Chicago (B.A. 1979). Wilfrid Sellars directed my doctoral dissertation at the University of Pittsburgh (Ph.D. 1984). The topic of this paper is Sellars' and Davidson's theories of meaning, which I will construe as comprising everything that they offer towards a theory of linguistic communication. However, as a witness to history I will allow myself some personal recollections.

I arrived in Chicago two years before Davidson, who was recruited after the closure of the philosophy department at Rockefeller University. I took his philosophy of language course in the Spring of 1977. In 1977-78 I was an exchange student at the University of Frankfurt, where Richard Rorty happened to be spending a semester as a guest professor. (Rorty told me that he had just sent off a book manuscript. That turned out to be *Philosophy and the Mirror of Nature* (1979).) Rorty revved up my enthusiasm for Davidson and, so, when I got back to Chicago, I asked him to direct my bachelor's thesis. I can remember the expression of reluctance on his face when he agreed, but he could not say no to my ebullience. The timing was not good. His wife was ill and he missed a lot of class. I met with him only about three times to discuss my thesis.

Davidson was a good teacher. He always came to class with something to say. We had assigned readings, but the lectures were more about ideas than about texts. With his radio-announcer's voice he always seemed to be on-air. I think the students tended to be afraid of him. Another undergraduate in the philosophy of language course confided to me that Davidson had written on his paper, "Try very hard to say something that you understand". But he had a nice side as well. One day I lost my mechanical pencil. Many weeks later I was in his office and found a mechanical pencil that looked just like mine lying on the edge of his

desk. I asked, “Is this my mechanical pencil?” He said that someone had left it there and he figured that whoever it was would sometime be back and claim it. I did not think he thought very much of me. Later Alan Donagan told me that Davidson had told him that if he had had the final version of my bachelor’s thesis before he wrote my letter of recommendation, he would have made sure I got into Princeton, which was some consolation.

As it was, I went to the University of Pittsburgh. That was probably better for me than Princeton would have been. At that time, in the early 80’s, Pitt was an exciting place for philosophy. (Brandom and Haugeland were assistant professors when I arrived, but I had already left before the arrival of McDowell.) In Frankfurt Rorty taught a course on Sellars. I did not take that course (I took two others from him), but nonetheless Rorty planted a seed of interest. So in Pittsburgh I threw myself in with Sellars. I took two courses from him for credit, including his course on the *Critique of Pure Reason*. After that I attended all of his courses, even his undergraduate courses, though I did not need them for credit. Towards the end of my time Sellars began to teach only his own views. He seemed to know some of his lines too well.

In our meetings to discuss my dissertation, we did talk and think together. I asked Sellars to read Tyler Burge’s “Individualism and the mental” (Burge, 1979), and we had many meetings in which he tried to come to grips with Burge’s social externalism. That was a good lesson for me in *focusing-in* on something, but there was much of my dissertation that we never discussed at all. Although he was the department’s “famous philosopher”, there were not many students working with him at that time. (Besides me, there were Pedro Amaral and Dionysios Anapolitanos. Bill de Vries had already finished.) I think he was grateful for my interest. I was tickled when at a department get-together he told people that if I didn’t exist he would have to invent me. Early in 1984 he had his stroke. He recovered enough to

return to teaching, but he was very frail and from that point on I worked on my dissertation alone.

Analytic philosophers are now wont to say that analytic philosophy does not now, if it ever did, analyze concepts and that analytic philosophy is distinguished from other brands of philosophy by its orientation toward problem-solving and careful argumentation. But in the 50's, 60's and 70's, analytic philosophers often did still seek to analyze concepts, and they might have done this with the presumption that putting forward explanatory hypotheses was not the proper business of philosophy. Sellars and Davidson have in common that they sought to analyze the concept of meaning. Moreover, both of them approached the analysis in terms of the concept of translation, although both of them might have denied that they did that. Davidson thought that he could explain linguistic communication in terms of his theory of meaning. Sellars thought that he could explain linguistic meaning without addressing interpersonal linguistic communication. Their analyses of meaning are in an interesting way complementary to each other, in that each provides something that the other fails to give, but both of them, I will argue, fall fundamentally short. I will contend that both of them were drawn away from the project of explaining the nature of linguistic communication by their focus on the analysis of meaning.

In my opinion, we do not need an analysis of the concept of meaning at all, but at most an explanation of the role that talk of meaning plays in linguistic exchange. Talk of meaning belongs, both in its everyday use and in its technical use in semantics and logic, to the metalinguistic discourse by means of which the speakers of a language normatively discipline their own language. This practice arises on top of a non-metalinguistic practice of using language to achieve practical ends, for it is that which talk of meaning seeks to discipline. The metalinguistic practice takes the prior non-metalinguistic practice as a given and attempts to improve it. In their explanations of linguistic meaning, Sellars and Davidson

neglected the nature of this more fundamental practice. They neglected, that is to say, the question of how, by means of language, people manage to cooperate with one another.

Davidson's theory of meaning

Davidson holds that a theory of meaning can take the form of a Tarski-style recursive definition of truth for a language and presents a methodology for discovering such a theory of meaning for a language. It is fair to characterize Davidson as holding that a theory of meaning for a language tells the possessors of the theory how to produce, for any utterance in the language to be interpreted, an utterance in their own language that expresses the same proposition. The only kind of linguistic communication that Davidson countenances is one that entails that interpreters attribute meanings to a speaker's words. Therein lies the proof, I will claim, that Davidson does not adequately address the nature of linguistic communication.

All students of philosophy who take a course in first-order logic are taught (I hope) how to write a finite definition, or finite list of axioms, from which can be derived infinitely many sentences of the form "*s* is true in *L* in model M if and only if *p*". On the left-hand side, in place of "*s*", will appear the quotation-name of one of the infinitely many sentences of *L*. M is a *model*; it contains a domain D of objects and an assignment Σ to each non-logical term of an appropriate extension drawing from the domain. For example, it might be that Σ("*a*") = $o_{27} \in$ D and that Σ("*F*") = $\{o_1, o_{27}, o_{106}\} \subseteq$ D. The clauses of the definition (or the axioms) will specify how the truth of each form of sentence depends on the extensions or truth-values of its immediate constituents. From such a definition one can derive, for each sentence of the language, a biconditional that expresses the truth conditions of that sentence in the given language relative to M. For example, the truth-conditions of an atomic sentence "*Fa*" will be specified thus:

"*Fa*" is true in *L* in M if and only if Σ("*a*") ∈ Σ("*F*").

The truth conditions for a sentence "$\forall x(Fx \rightarrow Gx)$" will be specified thus:

"$\forall x(Fx \rightarrow Gx)$" is true in *L* in M if and only if Σ("*F*") $\subseteq$ Σ("*G*").

(Here I suppress the intervening step of defining truth on a model relative to a *variable assignment*.)

Davidson supposes that this sort of theory of truth relative to a model for an artificial language can serve as a paradigm for a similar kind of theory of truth for a natural language, despite the fact that the syntax of a natural language will be much less transparent than that of the artificial languages of formal logic (TM). Davidson supposes (apparently, although he does not say this) that there is one special model for a natural language, call it *the intended model*, such that the truth of a sentence in the language on the intended model is identical to *truth* in that language (*simpliciter*). In the case of the intended model, we can take the specification of truth conditions generated by the definition of truth in a model and substitute a sentence of our own language that says the same thing. For example, where Σ^* is the assignment in the intended model for German, so that Σ^*("Fido") = Fido and Σ^*("ist ein Hund") = the set of all dogs, the theory of truth for German relative to the intended model should yield the following theorem:

"Fido ist ein Hund" is true in German if and only if Fido is a member of the set of all dogs.

For the right-hand side of such a specification of truth conditions we can substitute a necessarily equivalent sentence of ordinary English, yielding, in this example, the following conclusion:

"Fido ist ein Hund" is true in German if and only if Fido is a dog.

Sentences like this, of the form "*s* is true in *L* if and only if *p*", in which *p* translates the sentence that *s* names into the metalanguage are what Davidson calls *T-sentences*. In this way, Davidson supposes, we can generate for each sentence of a natural language a T-

sentence of this kind. A theory of truth for a language that yields in this way a T-sentence for each sentence of the language can double, says Davidson, as a theory of meaning for the language (TM 24).

It may be objected that we can state the conditions under which a sentence is true without yet specifying the meaning of the sentence (because two sentences with different meanings can be true under the same conditions), but Davidson holds that the theory of truth will suffice as a theory of meaning so long as it is the product of an adequate methodology (RI 139; RR 224). A simple account of the methodology, which he presents in several papers (TM 27; RI 134–135), goes like this: The interpreter observes the circumstances under which sentences of the language are held to be true. The interpreter may assume that the truth conditions of the sentence may be specified by specifying some circumstance that actually obtains on the occasion of utterance. This assumption is part of what is called *the principle of charity*. If a sentence is uttered on multiple occasions, the interpreter may be able to identify a circumstance that all of those occasions have in common and is warranted in hypothesizing that in general an utterance of the sentence is true if and only if that circumstance obtains. Such hypotheses form the data for a theory of truth. For instance, from the fact that it is generally raining at the time and place where speakers of German hold “Es regnet” to be true, we may infer that for every speaker and every time, “Es regnet” is true in German if and only if it is raining at that time in the vicinity of the speaker. The theory of truth for a language is constructed by attempting to construct a theory of truth from which as much of the data as possible can be derived. In order to produce a reasonably simple theory, we may be compelled to conclude that some of the hypotheses that we took as data were false and have to be discarded.

Before I go on, I have to take note of a lacuna in the account of the methodology that I set out in the previous paragraph. Davidson often writes as if the data for a theory of truth for

a language consists of T-sentences (TM 23; RI 133), for what can be derived from a theory of truth for a language is just such T-sentences. But the methodology calls on interpreters to find the circumstance that is common to many different *utterances* of the sentence that speakers of the language take to be true. Thus, one would expect the data ultimately to consist in hypotheses about what is usually the case when a certain form of sentence is *uttered* (as I said in my summary of the methodology). Consequently, we have to acknowledge a step in which we pass from facts about the circumstances in which sentences are uttered to general T-sentences identifying the conditions under which an utterance of a sentence is true for a given speaker at a given time (TM 34; RI 135).

Davidson seems to have supposed that the relativization to speaker and time is sufficient to capture the ways in which the truth of a sentence is relative to circumstances of utterance, but it is not. The truth of a sentence containing a quantifier (such as "Everyone is present") is relative to a domain of discourse. A sentence containing an incomplete predicate such as "ready" (as in "Tipper is ready") is relative to a specification of an activity (for which a thing is said to be ready). And so on. In order to capture the logical properties that such context-relativities induce, we need to characterize *contexts* as formal structures that provide a value for every contextual variable and then formulate our theory of truth as a theory of truth relative to arbitrary contexts (Gauker "What Tipper is Ready For"). Then we can say, for instance, that a necessary condition on the logical validity of an argument is that for every context in which the premises are true the conclusion is true. Once we have defined logical properties in terms of the class of possible contexts, nothing is added by introducing the speakers and times (actual or possible) whose speech acts they might characterize.[1]

[1] Davidson cannot easily accommodate this context-relativity without compromising his fundamental assumptions. Davidson holds that the only role played by the relation of reference (for example, between a singular term and an object or between a predicate and its extension) is in allowing us to write a finite theory from which all of the T-sentences for a language can be derived (RR). No

Davidson takes his theory of meaning to be at the same time a theory of linguistic communication. He believes that interlocutors, insofar as they understand what others say, can be interpreted as employing in effect a theory of meaning in his sense (NDE 256; SCT 311–312). He does not think that they have a theory of truth written out in their brains, but when an interpreter understands a declarative sentence uttered by a speaker, the interpreter can represent in his or her own language the conditions under which the speaker's sentence is true (SCT 312). Davidson denies that a theory of truth is a theory of translation, but his objection to that construal is just that we could have a method of translating one language into another though we did not understand either the language being translated or the language of the translation (RI 129). No such objection arises if we say that a theory of meaning, on Davidson's account, enables an interpreter to find, for each utterance of a sentence in the object language, a sentence of his or her own language that, in the context of the interpreter, expresses the same proposition.[2]

To the theory so construed, one might object that we cannot always find in our own language a sentence that, in our own context, has the same meaning as the sentence that we want to understand. When using a foreign language, one seldom finds an exact synonym in one's own language for words of the foreign language. The best explanation of the meaning

---

reference relation needs to be countenanced in formulating the evidence for such a theory. It is a fundamental principle for him that the evidence for a theory of truth for a language consists of statements about the conditions under which speakers of the language hold sentences to be true. However, in view of context-relativity, it is inevitable that we will formulate such statements in terms of reference. For instance, we will have to say things such as the following: Speakers of German hold the sentence "Du bist krank" to be true in a context if and only the person that "du" *refers* to in that context is in the *extension* that "krank" has in that context. Thus, relations of reference must be mentioned even in stating the evidence for the theory.

[2] Here I follow Kaplanian orthodoxy (Kaplan 1989) in supposing that while *sentences* have *meanings*, *utterances* express *propositions*, although Davidson himself does not make use of this distinction.

of a foreign word will often convey understanding by providing many examples of the use of the foreign word together with a translation suitable for that example. Non-native speakers contemplating a use of a foreign word can only check to see whether the contemplated use is close enough to one of the examples provided. If in German a speaker says, regarding a particular family, that praying before meals is their "Gewohnheit", is the speaker saying that praying before meals is a *habit* of theirs or that it is a *custom* of theirs? That depends on what exactly it is that they do. Shall we translate the speaker's words as meaning that praying before meals was "either a habit or a custom"? Well, no, because for that particular utterance there may be only one correct choice. In German, there are many words for revolt: "Aufstand", "Aufruhr", "Aufbegehren", "Auflehnung", "Erhebung", "Revolte", "Rebellion". If we want to say in German that the slaves in a mine staged a revolt, which of these is the correct translation of "revolt"? It's the same with verbs. The German verb "gehen" can mean just *go*, allowing various modes of transport, or specifically *walk*, for example, in the sentence, "Wir gehen ins Restaurant". Even if on some particular occasion it is clear which translation is best, the decision procedure will not involve the application of precise criteria but will require instead all-things-considered judgments regarding similarities to paradigms. Translations generated in this way cannot in addition be generated by a derivation from a theory of truth modeled on the model theory of formal logic in the way Davidson imagines.

For these reasons I consider it a foregone conclusion that we will not succeed in producing in English a theory of truth for another natural language of the kind that Davidson envisions. An answer to this objection might be that the metalanguage of the theory of truth is not any spoken language but rather a *language of thought*. Since this language of thought is the medium for every thought that we might in any language wish to express, there must be, for every utterance of a sentence of any spoken language, a sentence in the language of thought whose content in context matches the meaning of that utterance. I believe that the

language of thought hypothesis has been amply refuted (Gauker *Words and Images*, ch. 4), and there are aspects of Davidson's philosophy that strongly suggest that he would reject it too (TT). But if we suppose that Davidson accepts it, then we can set aside a fundamental objection in order to evaluate the rest of what Davidson has to say, which is what I propose to do.

A theory of the truth conditions for sentences of an object-language, as formulated in a metalanguage, provides the meanings of the sentences of the object language, in a manner that serves communication, only on the assumption that the interpreter understands the sentences of the metalanguage. We still need, it may be objected, an account of what each interpreter understands in understanding the sentences of the metalanguage. This objection gives us another reason to interpret Davidson as supposing that there is a language of thought. If there is a language of thought, then one can use it in thinking without having to interpret the sentences in it by means of which one thinks. When it comes to the language of thought there is no question of how to interpret it in one's own case, but only the question how to tell what other people are thinking. The only question that remains is how to figure out what people believe and desire and, in general, what they *think*. One might suppose that there is, in addition, the question: What does a thought's having its content consist in? I am not sure that Davidson ever countenanced that question at all.

I have just sketched one route to the conclusion that Davidson's theory of the interpretation of a speaker's words must be supplemented with a theory of an interpreter's beliefs. A route to this same conclusion that lies more on the surface of Davidson's writings begins with the observation that plausible hypotheses about truth conditions go hand-in-hand with plausible interpretations of speakers' mental states (see esp. BBM). On Davidson's account, the more encompassing theory of interpretation entails a considerable complication of the methodology. We have to make sure that our ascriptions of belief and intention show

that the agent conforms more or less to the prescriptions of decision-theory (BBM 147–148; SCT 316–318), that our interpretations respect the evidential relations countenanced by the agent (SCT 322) and that the agent's perceptual beliefs (beliefs formed on the basis of perception) are grounded in perception in the right way (SCT 320). These additional constraints on the ascription of belief and intention to an agent have a bearing also on our account of the truth-conditions of the sentences of a language, because our hypotheses about the meaning of the agent's words will depend on what we take the agent to believe and intend (BBM 147; SCT 322).

If in this way, namely, by citing Davidson's more general theory of interpretation, we answer the objection that Davidson needs a theory of what an interpreter understands in understanding the meaning of a sentence, then we are still left with an incongruity between what Davidson says about linguistic meaning and what he says about mental content. There is supposed to be a particular form that a theory of linguistic meaning assumes, that of a definition of truth. This provides us with a distinction between what the theory says and the methodology for discovering the theory, as well as a rather clear account of the latter. Davidson does not likewise insist on any particular form for his theory of the beliefs and intentions of an agent, and the methodology for attributing beliefs and intentions amounts to hardly more than "Try to make sense of everything he or she does, all things considered". If, as I suppose, Davidson's answer to the objection that we still need an account of what a thinker understands in understanding the meaning of a sentence is supposed to be his general method for intentional state attribution, then his answer turns out to be rather thin.

A peculiarity of Davidson's theory of communication is that he sees no necessity in a common language. In principle each person could have his or her own language, spoken by no one else, provided only that each interlocutor possessed an adequate theory of truth for the language of every interlocutor (CC 276; NDE *passim*). We could be a shipload of pirates,

each of whom speaks his or her own personal language, one a variant of English, one a variant of Chinese, one a variant of Arabic, and so on, provided only that the English-speaker could specify in his or her own language the truth conditions of the sentences of the Chinese-speaker, and *mutatis mutandis* for each pair of pirates. I contend that Davidson's pirate-ship model of linguistic communication is untenable because it completely rejects the social character of linguistic meaning. As Putnam (1975) and Burge (1979) taught us, the meaning of the words in one speaker's mouth cannot be identified apart from the meaning of those words in the mouths of other members of the community. Meanings will have this social character in every community that reaps the benefits of what Putnam calls the division of linguistic labor. (This social character need not be limited to cases in which a subcommunity qualifies as the experts about the subject matter that the words are used to describe. See Gauker 2007.) In various places Davidson addresses the social externalisms of Burge and Putnam. He accepts from Putnam the relativity of content to physical environment (KOM 39). But he steadfastly insists on the availability of non-social, individualistic interpretations of words and the thoughts that words express (EE 197–199; SAL 120–121; KOM *passim*).[3]

The closest Davidson comes to endorsing a model of linguistic cooperation is his endorsement of Grice's implausible analysis of meaning something by something (CC 277).[4]

---

[3]The claim that Davidson ignores the social character of meaning might seem surprising in light of his notorious thesis in "Thought and Talk" that only creatures who belong to a speech community can have beliefs. (TT 170). But Davidson explicitly says there (TT 157) that he does not mean that one has to speak the same language as another in order to have beliefs, only that one must be able to interpret what other people say. There seems to be an uncharacteristic concession to Burge in MS 49.

[4] I call Grice's theory implausible because he assumes that speakers have extremely complex intentions of a kind that in fact no human being in history has ever actually had. In the quotation that follows, Davidson paraphrases only the simpler account in Grice, "Meaning", but if one accepts Grice's arguments for that, then one should accept as well his arguments for the much more complex analysis in Grice, "Utterer's Meaning and Intentions". For my critique of Grice, see Gauker *Words without Meaning*.

> What matters to successful linguistic communication is the intention of the speaker to be interpreted in a certain way, on the one hand, and the actual interpretation of the speaker's words through the interpreter's recognition of the speaker's intention, on the other (SCT 311).

Davidson cannot acknowledge the social character of meaning, because doing so would be incompatible with the Gricean model of communication as well as with any other model that holds that communication consists in a speaker's using a sentence to reveal an underlying thought and the hearer's recognizing the thought that the speaker reveals. In the context of such a theory, the social character of meaning can be developed in two ways. One possibility is to say that the content of the thought revealed is distinct from the content of the utterance of the sentence spoken. In that case, however, we will have to say that thoughts will not be effectively expressible in public language at all, due to the necessary difference in content. The other possibility is to say that the content of the thought revealed derives its content from the content of the sentence used to express it and is therefore also social in character. In this case, there is no reason to mention the thought in our account of literal, undisguised speech. We can say simply that communication consists in a speaker's making an utterance with a certain content and the hearer's recognizing that the speaker's utterance has that content.[5] Further, acknowledging the social character of meaning will imply that Davidson's list of constraints on interpretation is critically incomplete. What it leaves out are the norms governing interpersonal cooperation by means of a shared language.

We have seen that Davidson says that the interpretation of another person's words requires the application of a theory of meaning. This contention can be derived from his other

[5] In Gauker, "Social Externalism and Linguistic Communication", I argued in a different way that the social externalism about content cannot be combined with what I there called the expressive theory of communication.

premises. From the pirate-ship model it follows that there is no linguistic communication without interpretation. Since no two interlocutors speak the same language, each interlocutor is in the position of having to interpret the words of everyone else. From the Gricean model, it follows that there is no interpretation apart from understanding the speaker's communicative intention. If we add that the recognition of the speaker's communicative intention requires the application of a theory of meaning, which Davidson also says (SCT 312), then we may conclude that interpretation requires the application of a theory of meaning. When the absurdity of the assumption that we are constantly applying a theory of meaning in understanding the words of another is pointed out to him, Davidson defends himself only by protesting that the process need not be conscious (SAL 112, 122).

Whether or not we call it *interpretation*, there is a basic kind of linguistic communication that does not require an interlocutor to apply theories of meaning to the words of the other interlocutors. Interlocutors, I contend, may have the ability to engage in the basic practices of linguistic communication without giving any thought to the meanings of words, whether conscious or unconscious. They can engage in routine linguistic exchanges of a kind that give language an important function in facilitating cooperation in the achievement of practical goals without ascribing meanings to the words that others speak. Consequently, they also do not always have the communicative intentions that Davidson insists they must always have. I will present an example in the next section. If he had said that there is a kind of linguistic communication that does not require interpretation, then Davidson might have held that all *interpretation* requires the application of a theory of meaning while acknowledging in addition a kind of linguistic communication that does not require interpretation. But due to his focus on interpretation, construed as requiring the application of a theory of meaning, this more basic kind of linguistic communication does not catch Davidson's attention at all.

Sellars' theory of meaning

Unlike Davidson, for whom the theory of meaning is hardly more than an account of how to attribute truth conditions, Sellars goes out of his way to deny that the essence of a meaning lies in being grasped. The meaning of an expression (word, phrase or sentence) is its functional role in a language game, and the functional role of an expression is defined by three kinds of "move" or "transition" (SRLG §§21–23, 328–329; SM 114; MFC 423). One kind is *language-entry transitions*, which concern verbal responses to non-verbal circumstances. The second kind is *intra-linguistic moves*, which concern the passage from one linguistic utterance to another. The third kind is *language-departure transitions*, by which an agent passes from a linguistic utterance to a nonlinguistic undertaking. The totality of such rule-governed transitions pertaining to an expression is supposed to implicitly define the meaning of the expression, in the sense of providing the basis for anything we might like to say about the meaning of the expression (SRLG §31, 332). In Sellars' focus on rule-governed linguistic moves and transitions, he neglects, I claim, a more fundamental question, namely, concerning the basis for the moves and transitions, and thereby neglects as well the nature of linguistic communication. Moreover, he assumes that the moves and transitions that define the meaning of an expression pertain specifically to a nondialogical, *intra*personal use of language, although he repudiates the conception of communication that would have forced him to do so. In this way too he fails to come to grips with the way language serves interpersonal communication.

The starting point for Sellars' reflections on meaning in "Some Reflections on Language Games" (SRLG) is the question of what it means to follow a rule. Since we want to say that speaking a language involves following a rule, we do not want to say that following a rule entails understanding a verbal formulation of the rule. On the other hand, we want to say

that following a rule requires more than acting in conformity to it. Sellars' solution is to say that an organism may, in a sense, be designed to follow the rule, either because it is the product of evolution through natural selection, or because its behavior is the product of a kind of learning that somehow ensures conformity to a whole system of rules. In order to use an expression meaningfully, according to Sellars, it is not necessary to have any kind of metacognitive awareness of the rules that govern the moves and transitions; one must just be able to act in a way explicable as due to the whole system of moves and transitions. One may equally think of these rules governing moves and transitions as rules of thought, because they can be followed even without the possession of a prior ability to conceptualize the circumstances under which they apply. Sellars says that they are *ought-to-be*'s, not *ought-to-do*'s (LTC 510; SM 76). In SRLG, Sellars says that acting in conformity with language-entry and language-departure transitions cannot rise to obedience to a rule (§§35–36, 334), but in later writings he referred to all three kinds of move and transition either as themselves rules (SM 114) or as rule-governed (MFC 423). It is correct to describe the moves and transitions as rules or as ruled-governed, and not merely as regularities, because they have a normative force. A person can be criticized for not conforming to them (SM 157).

Sellars does not quite propose to appeal to these moves and transitions in *defining* the meaning of any given expression (word, phrase or sentence). Rather, he appeals to them in explaining what it means to say of some expression that it has a certain meaning (which is not quite the same thing as giving a definition). If we say something of the form "*E* means *m*", then *E* is the name of an expression, the meaning of which is to be explained, and *m* is an expression of the metalanguage, which is here used, not in the normal way, but, rather, as a label for a kind of linguistic role-player (SRLG §31, 332; MFC 431). "*m*" here denotes a *role* that the expression "*m*" normally plays in the metalanguage, and that role is supposed to be specified by the totality of moves and transitions of the sort described above. One can say

that for Sellars meanings are expressed by translations, inasmuch as all he proposes to say about meanings *per se*, is that sentences of the form "*E* means *m*", which in effect provide translations, are to be explicated in terms of the rule-governed moves and transitions. It is important that "*m*" in this formula be an expression that those ascribing meaning to *E* understand, because in that case the meaning statement can succeed even in the absence of an account of the moves and transitions governing *E* (SM 130).

Sellars's theory of meaning faces two fundamental problems, which I will call the *empirical problem* and the *metaphysical problem*. The empirical problem is that it is hard to think of enough good examples of the sorts of moves and transitions that Sellars needs in order to be confident that the totality of such rules will distinguish the meaning of each expression from the meaning of each other expression. There are seldom any linguistic utterances that really *ought to be* in Sellars' sense. Sellars remarks in places that the rules include prohibitions (MFC 422) and may include permissions as well as imperatives (LTC 512, note 4), but it is also hard to think of any rules of these kinds that might play a role in determining meaning. What kind of intra-linguistic moves play a role in determining the meaning of the word "dog"? Yes, it is a rule that from "*x* is a dog" one may infer "*x* is an animal". But moves like this, which express analytic necessities, will not take us very far in distinguishing the meaning of "dog" from the meaning of "cat". Is there a rule-governed move from "*x* is a friendly, furry animal that makes yelping and growling noises" to "*x* is a dog"? Maybe there are some friendly foxes that make yelping noises.

It is obvious that the moves and transitions will distinguishing the meanings of words only if they include also language-entry transitions and language-exit transitions. What kind of rule-governed language-entry transitions play a role in determining the meaning of "dog"? Is there a rule that says that whenever one is looking at a dog in sunlight one may declare "That is a dog"? Is that so, even if the dog looks like a sheep and one gets only a fleeting

glimpse of it? What kind of rule-governed language-exit transitions are there for "dog"? Is there a rule that says that if someone is in the mood to pet a dog and says to him- or herself, "This is a dog", then he or she must (or may) thereupon pet a dog (compare SRLG §66)? It is just completely unclear what kinds of rule-governed moves and transitions are supposed to do the job of determining meaning.[6]

The metaphysical problem is that, even if we were given a large collection of plausible-sounding rules, it might still be quite unclear how conformity to them would ensure that the language served the purpose of interpersonal cooperation. Suppose that the rule-governed language-entry transitions ensure that entry-level linguistic responses usefully classify objects into kinds. So the rules governing "That's a dog," "That's a horse" and That's a pig", etc., ensure that words like "dog", "horse" and "pig" are used to draw a useful distinction between kinds of animals. But this fact about language-entry transitions will not all by itself ensure that the language serves interpersonal communication. We can invent a collection of rules that take "That's a dog," "That's a horse" and "That's a pig", and so on, as inputs, and then, in each case, instruct us to infer only "How very nice!". In this case the useful classification that the language-entry transitions initiated will be rendered inert. The useful work done by the rules in one region of the body of rules is always at risk of being undone by the rules in another region.

What this shows is that not only the language-entry transitions but also the totality of the intra-linguistic moves (though perhaps not each individual intra-linguistic move) must somehow facilitate interpersonal cooperation. It is not clear in advance what would count as

[6] There is a narrower program in the philosophy of logic that aims to explain the meanings of logical connectives in terms of the inference rules of a formal deductive calculus (Peregrin) or tries to show that the inference rules for a connective determine the function over extensions that the connective represents in a model-theoretic semantics (Murzi). The empirical problem for the more general inferentialist program, which I have identified here, does not touch this narrower program.

explaining how the totality of rule-governed moves and transitions facilitates cooperation. One does not *explain how* they facilitate cooperation just by saying, "Well, we follow the rules and, see, it results in cooperation". It is plausible that we would need to show that the totality of rules somehow articulates the structure of reality in a way that is useful for human purposes.[7] One could still maintain that it is only the rule-governed moves and transitions that we need to appeal to in explaining meanings. Nonetheless, we would need this deeper explanation of the rationale behind the rules in order to understand how language functions to facilitate interpersonal cooperation.

Sellars very clearly countenances this objection in SRLG, when he writes that "the pragmatist"

> argues that to conceive of a language as a game in which linguistic counters are manipulated according to a certain syntax is to run the danger of overlooking an essential feature of languages — that they enable language users to find their way around in the world and satisfy their needs (SRLG §47, 340).

Well said! Unfortunately, Sellars offers no satisfactory answer. What he says in response is only that it is a category mistake to hold that the meaning of a sentence should be explicated in terms its role in problem solving behavior (SRLG §49, 340) and that the rules of language codify material inference moves and govern also language relating to conduct (SRLG §50, 340–341).

[7] Indeed Sellars contended that his conception of language as governed by functional rules could be combined with a conception of certain sorts of linguistic utterances as constituting pictures of reality (SM 136). How these two different approaches to the nature of language can be coherently combined is a notoriously difficult question for Sellars scholarship (Harman 418–419 ; Rorty 295–299; deVries 250–256).

Sellars seems to assume that the moves and transitions that define the meanings of words are only such as might be extracted from a purely *intrapersonal* use of language (even if they also govern some interpersonal exchanges). His examples, such as they are, pertain only to a person's linguistic response to the world, to his or her further processing of those responses and to his or use of language in guiding his or her own actions (SLRG §§22, 23, 329). The pertinent intra-linguistic moves are exclusively inferences, not contributions to dialogues such as take place between two people. Moreover, it is only such intrapersonal exercises of language that Jones, in the Myth of Jones, takes as his model for thought. When Jones discovers the existence of thoughts, he starts by observing conduct "threaded on a string of overt verbal episodes". Sellars (but not Jones) calls these overt verbal episodes collectively "thinking out loud" (EPM §56, 186–187; SM 75). Jones in effect discovers thoughts when he conceives of silent acts of inner speech that accompany conduct when it is not accompanied by thinkings-out-loud. These acts of inner speech are conceived of as analogous to, and are modeled on, thinkings-out-loud.

These acts of inner speech in terms of which Jones explains the behavior of his fellow Ryleans are episodes of nondialogic, self-directed speech. For instance, a hungry person is supposed to say in inner speech, "Here is an edible object" and thereupon proceed to eat it (EPM §56, 186–187). In an example from class, which Sellars used to illustrate the kind of thinking out loud that serves as our model for occurrent thoughts, he asked the students to imagine him arriving at a mailbox intending to mail some letters and then striking his forehead with his palm and declaring, to no one but himself, "I forgot to bring the letters!" In LTC Sellars explicitly says that the thinkings-out-loud on which thoughts are modeled are not directed at an audience and are not illocutionary performances such as statements, and he warns the reader not to treat linguistic communication as the model for thought (LTC 517–518; see also MTC 420; SM 95, 112n1, 157).

That Sellars explicates meaning exclusively in terms of nondialogic self-directed speech should be puzzling, because Sellars goes out of his way to reject a conception of other-directed speech that would prevent him from using other-directed speech as his model for thought. If it were the function of language to reveal to other people the content of underlying thoughts, then it would not make sense to treat interpersonal speech as the model for thought. The thoughts modeled on speech would in turn have to be expressions of some other, underlying thoughts, which makes no sense (cf. LTC 510). But Sellars explicitly denies that spoken language should be conceived as a means of expressing thoughts in this way.

> Although [Jones'] theory postulates that overt discourse is the culmination of a process which begins with "inner discourse," this should not be taken to mean that overt discourse stands to "inner discourse" *as voluntary movements and intentions stand to intentions and motives*. True, overt linguistic events *can* be produced as means to ends. But serious errors creep into the interpretation of both language and thought if one interprets the idea that overt linguistic episode *express* thoughts, on the model of the use of an instrument (EPM §58, 188).

(No other short passage of philosophy has had more influence on my own philosophical development than this one.) On the contrary, says Sellars in the very next sentence, the capacity to speak out loud is so far independent of the capacity to think in silence that the latter is acquired only by means of first acquiring the former (an idea on which Sellars elaborates in LTC 521–522). In this passage, I take it, Sellars rejects the model of linguistic communication according to which a speaker chooses words that reveal the content of an underlying, language-independent thought and the hearer recognizes the content of the thought that underlie the speaker's words.

Accordingly, it would have been quite open to Sellars to treat Jones as modeling thought on the kind of interpersonal discourse in which two people exchange information and jointly plan a course of action that serves both their ends. In other words, the category of thinkings-out-loud might have included episodes of dialogic interpersonal speech. We need not assume that routine verbal exchanges between people involve each party's forming an intention to produce a certain result in the hearer, such as leading the hearer to recognize an underlying thought in the speaker, and then acting in such a way as to realize this intention. Most of the verbal exchanges that occur in our lives involve no such complex intentions. If a tourist stops me in the street and asks me, "Where is the St. Peter's church?", I reply without deliberation, thus: "Go to the end of this street. Turn left in front of the crucifix. Walk until you reach an arch on your right. Go through that arch. You will see it". Our speech acts in such cases may be intentional in the weak sense that we remain in control and can alter course if something contrary to expectation happens. The tourist designs a question, and I design an answer, but our intentions do not go any deeper than that. I need not think about the tourist's intentions in asking. The tourist need not contemplate my motives in answering as I do. Just so, one can ask oneself a question silently in thought and answer it silently in thought as well.

Moreover, I think it would clearly have been better for Sellars to suppose that episodes of interpersonal dialogue serve as our model for intrapersonal thought. The reason is that episodes of self-directed talk serve no real purpose. There is really no reason for Jones' companion to say to himself "Here is an edible object" before he proceeds to eat it. But if Jones' companion is talking to someone who is hungry and who does not recognize that the object before them is edible, then there is reason for the agent to say this; saying it produces a useful result. Likewise, in Sellars' story about the mailbox, nothing is really achieved by his saying out loud, "I forgot to mail the letters". In order to understand how thinking produces

useful results, we should model thinking on the kind of speech that produces useful results, which is dialogue between different people. Self-directed speech out loud can serve a useful purpose insofar as it is itself a kind of internalized (but not silent) interpersonal dialogue, but we might as well take full-fledged interpersonal dialogue as our model instead of restricting our attention to the overt speech that is in turn an internalization of interpersonal dialogue.

In sum, while Sellars might have told a story in which Jones discovers thoughts by positing internal, intrapersonal conversations analogous to overt, interpersonal conversations, he does not tell that story but instead tells a story in which Jones models thoughts on nondialogic self-directed talk. Why? Here is my hypothesis: While Sellars does reject the model of communication on which speakers reveal the contents of language-independent thought to hearers, he nonetheless thinks that interpersonal discourse requires the application of rules he calls *ought-to-do*'s. The teachers of a language (at least they) must be able to apply ought-to-do's (LTC 512). An ought-to-do rule can be followed only by an agent who has the capacity to conceptually represent the conditions under which it applies (LTC 509). It is for that reason that interpersonal discourse is not suitable as a model for thought. We would create a regress if we modeled thinking on moves and transitions in language and then said that the moves and transitions that take place in thought are possible only by means of thoughts representing the conditions under which they may take place. My disagreement with Sellars concerns precisely this point. I see no reason to think that routine interpersonal exchange, though it does involve asking questions and answering them with statements, requires the interlocutors to conceptually represent the conditions under which linguistic rules are applicable. So, from my point of view, routine interpersonal discourse can serve perfectly well as a model for intrapersonal thought.

Even if Sellars had not had this reason for confining thinkings-out-loud to nondialogic self-directed speech, he would have had another reason to do so. Sellars is committed to

thinking of the uses of language that serve as a model for thought as governed by rules. He is committed to that, because he is committed to explaining meaning in terms of such rules. It is easier to suppose that rules of some kind govern nondialogical self-directed speech than it is to suppose that there are rules governing interpersonal speech. While the empirical problem of identifying the rules governing the moves and transitions in terms of which meanings are to be explicated is already very hard on the assumption that the rules govern exclusively intrapersonal discourse, it is even harder if the rules are supposed to be rules governing interpersonal discourse. If a speaker says to a hearer, "There is firewood stacked next to the white boulder", what rule determines what the hearer must or may do in response? Between a speaker's utterance and a hearer's response there intervenes the hearer's deliberation, which undermines any attempt to state a general rule. In this respect too, I conclude, Sellars' focus on the analysis of meaning led him away from a deeper examination of the functionality of language.

Lessons learned

Sellars' theory of meaning is in one respect superior to Davidson's in that he does not beg the question. Davidson begs the question in not offering theories of mental content comparable to the theories of linguistic meaning. Sellars is not subject to this objection because he reduces meanings to rule-governed moves and transitions. He does not appeal to thoughts as the medium of understanding meanings, and he is prepared to say the same thing about thoughts that he says about meanings, on the grounds that thinking (by means of concepts) is just an internal form of speaking in a language.

Davidson's theory of meaning is, on the other hand, in one respect superior to Sellars'. Precisely because he reduces one's understanding of a language to a translation into the medium of one's own thoughts, he is not driven to posit a set of rules of use that we are

unable to identify and which, if we could identify them, we would be unable to justify. We have an ability to think—it is not our philosophical business to explain how it works—and we understand another language insofar as we understand how to translate its sentences into our thoughts.

The objections I have raised against both authors can be summed up as follows: Both authors, through their focus on a theory of meaning, are led to neglect the nature of interpersonal linguistic communication. In the case of Davidson, it is his assumption that interpretation requires the attributions of meanings that leads to this neglect. In the case of Sellars, it is his assumption that interpersonal communication rests on an ability to conceptualize, as well as his assumption that meaning must be definable in terms of rule-governed moves and transitions that leads to this neglect.

What might a theory of meaning look like if it did not try to embed in it a theory of linguistic communication? What might a theory of linguistic communication look like if it did not have to take it to have the form of a theory of meaning? Here I will briefly sketch the answers I have developed in more detail in other publications (Gauker "The Circle of Deference", *Words and Images*, "The Illusion of Semantic Reference").

A theory of linguistic communication must begin with *basic* linguistic communication. I define this as communication by means of language in which one person informs another person about features of the environment in a manner that will enable the interlocutors collectively to achieve their individual or shared goals. I assume, as I have already illustrated, that such behavior does not take the form of interlocutors revealing to one another the contents of underlying language-independent thoughts. Since it is not that, we are free to model elementary processes of *intra*personal thought on such interpersonal discourse. We may think of one kind of thinking as a kind of inner dialogue in the very languages we speak (Gauker *Words and Images*, "Inner Speech").

Of course, all speech is itself the product of some kind of thinking. Some kind of thinking goes into learning how to use a language, and some kind of thinking goes into choosing words to speak and responding to words that are spoken to one. So there has to be some kind of thinking other than thinking by means of the language one speaks. My hypothesis, developed in depth in my 2011 book (Gauker *Words and Images*), is that we can explain basic linguistic communication as the product of *imagistic cognition*. In imagistic cognition, we solve problems by means of manipulating mental images (which need not be conscious) (Gauker "Amodal Completion"). For instance, I use imagistic cognition in taking apart a faucet and putting it back together again, because I use mental imagery in recognizing that two parts will fit together before I try to fit them together. We may choose our words on the basis of the mental imagery that we engage in. Hearing the words that other people speak can have an effect on the courses of mental imagery that hearers engage in and thereby affect what hearers do, roughly, by confining hearers to courses of mental imagery that would elicit the same speech that the speaker produces.

The picture is expressly *not* that the speaker conveys to a hearer his or her mental images. First of all, that is not true. If something I say to you is elicited by some mental imagery of mine, I cannot expect that upon hearing my words you will form images of the same kind. Second, mental images do not have the kind of content that the utterance of a sentence might be thought to have. A mental image of a dog does not classify the thing that it represents as a dog. The effect of speaking on the mental imagery of a hearer is much more indirect than this. Nonetheless, my contention is that mental imagery can elicit speech and that the speech elicited can bring about cooperative activity by means of the constraints it places on other people's mental imagery.

Much of linguistic communication is not *basic* linguistic communication in this sense. Where a practice of basic linguistic communication is in place, episodes of using language

can become the subject of further linguistic discourse. Unobservable episodes of linguistic discourse, called *thoughts*, can also become the subjects of further discourse. (These may be unconscious for the agents engaged in them.) It is in this context that a more advanced kind of discourse becomes possible, in which people choose their words in light of their thoughts about other people's thoughts and hearers respond to what people say in light of their thoughts about what the speaker might be thinking.

The theory of *basic* of linguistic communication, and much of the theory of more advanced linguistic communication, can be developed without saying anything about *meaning*. What we need to explain about meaning, I contend, is not what meaning *is* but what *talk* of meaning contributes to linguistic communication. What talk of meaning contributes to linguistic communication is various devices for normatively regulating our language. We cannot expect to explain how linguistic communication works by explaining the meanings of linguistic expressions (words, phrases and sentences). But where a basic practice of using language to facilitate cooperation is up and running, talking of meaning can serve to grease the wheels and even to introduce permanent innovations that enable the machine to do new things.

We talk of meaning to grease the wheels when we explain what individual expressions mean. For example if A tells B that C was merely "dissembling" when C professed to have no romantic interest in D, and B does not know how to respond to this, but only because "dissembling" is not a word in B's working vocabulary, then A or anyone else can help B figure out how to respond by explaining that to say that someone is dissembling is to say that he or she is acting in a way designed to conceal his or her own feelings or thoughts without outright lying. We can call this explanation an explanation of the *meaning* of "dissembling" even though in giving the explanation we may not use the words "meaning" or "means" at all. One looks for and receives explanations of meaning of this kind very frequently in the

course of using a foreign language. Explanations of meaning of this kind are best viewed as *proposals*. In explaining meanings, one proposes to use words in a certain way. Precedent can be cited as justification for a proposal, but proposals are not strictly bound by precedents. Explanations of meaning in this sense are normative, but only in the sense that accepting the proposal that an explanation of meaning offers generates a normative commitment to conform to the proposal.

Beyond these everyday ways of talking of meaning, we can also talk of meaning in the context of what we might call the science of meaning. This is what philosophers, linguists and logicians are engaged in when they try to explain in precise ways the meanings of sentences. For example, one can try to define the meanings of natural language sentences in such a way that, given these explanations, one can literally prove that an argument consisting of natural language sentences is logically valid or not, as the case may be. This is the kind of project that I am engaged in, for instance, when I try to define the conditions under which conditional sentences are acceptable (or assertible) relative to a context (Gauker *Conditionals in Context*, Punčochář and Gauker). These efforts do not simply explain the given nature of the meanings of sentences as we found them. They inevitably involve a certain amount of regimentation amounting to a kind of recommendation. The hope is that when the work is well done the language regimented in this way will actually be a more useful tool for interpersonal cooperation.

Funding

This research was supported by a grant from the Austrian Science Fund (FWF) (Grant DOI: 10.55776/I6219; Project title: How can we think in a language?).

Bibliography

Burge, Tyler. “Individualism and the Mental”. *Midwest Studies in Philosophy, vol. 4, Studies in Metaphysics*, edited by Peter A. French, Theodore E. Uehling, Jr., and Howard K. Wettstein. University of Minnesota Press, 1979, pp.73–121.

Davidson, Donald. “A Nice Derangement of Epitaphs”. *Truth, Language and History*, Oxford University Press, 2005, pp. 89–107. Originally published in Grandy, Richard E. and Richard Warner, editors. *Philosophical Grounds of Rationality*. Oxford University Press, 1986, pp. 157–174. [NDE]

Davidson, Donald. “Beliefs and the Basis of Meaning”. *Inquiries into Truth and Interpretation*. Oxford University Press, 1984, pp. 141–154. Originally published in: *Synthese*, vol. 27, 1974, pp. 309–323. [BBM]

Davidson, Donald. “Communication and Convention”. *Inquiries into Truth and Interpretation*. Oxford University Press,1984, pp. 265–280. Originally published in: Dascal, Marcelo and Hubert Cuyckens, editors. *Dialogue: An Interdisciplinary Approach*. John Benjamins, 1985, pp. 11–26. [CC]

Davidson, Donald. “Epistemology Externalized”. *Subjective, Intersubjective, Objective*. Oxford University Press, 2001, pp. 193–204. Originally published in: *Dialectica*, vol. 45, 1991, pp. 191–202. [EE]

Davidson, Donald. “Knowing One’s Own Mind”. *Subjective, Intersubjective, Objective*. Oxford University Press, 2001, pp. 15–38. Originally published in: *Proceedings and Addresses of the American Philosophical Association*, vol. 60, no. 3, 1987, 441–458. [KOM]

Davidson, Donald. “The Myth of the Subjective”. *Subjective, Intersubjective, Objective*. Oxford University Press, 2001, pp. 39–52. Originally published in: Benedikt, Michael

and Rudolf Berger, editors. *Bewusstsein, Sprache und die Kunst*. Verlag der Österreichischen Staatsdruckerei, 1988, pp. 45–54. [MS]

Davidson, Donald. "The Social Aspect of Language" in *Truth, Language and History*, Oxford University Press, 2005, pp. 109–125. Originally published in: McGuiness, Brian and Gianluigi Oliveri, editors. *The Philosophy of Michael Dummett*. Kluwer Academic Publishers, 1994, pp. 1–16. [SAL]

Davidson, Donald. "The Structure and Content of Truth". *Journal of Philosophy*, vol. 87, 1990, pp. 279–328. [SCT]

Davidson, Donald. "Truth and Meaning". *Inquiries into Truth and Interpretation*. Oxford University Press,1984, pp. 17–36. Originally published in *Synthese*, vol. 17, 1967, pp. 304–323. [TM]

Davidson, Donald. "Radical Interpretation". *Inquiries into Truth and Interpretation*. Oxford University Press, 1984, pp. 124–139. Originally published in *Dialectica*, vol. 27, 1973, pp. 313–328. [RI]

Davidson, Donald. "Reality without Reference". *Inquiries into Truth and Interpretation*. Oxford University Press, 1984, pp. 215–225. Originally published in *Dialectica*, vol. 31, 1977, pp. 247–253. [RR]

Davidson, Donald. "Thought and Talk". *Inquiries into Truth and Interpretation*. Oxford University Press, 1984, pp. 155–170. Originally published in: Guttenplan, Samuel, editor. *Mind and Language*. Oxford University Press, 1975, Pp. 7–23. [TT]

deVries, Willem. *Wilfrid Sellars*. McGill-Queen's University Press, 2005.

Gauker, Christopher. *Words without Meaning*. MIT Press, 2003.

Gauker, Christopher. "Social Externalism and Linguistic Communication". *Meaning, Basic Self-Knowledge and Mind: Essays on Tyler Burge*, edited by Maria J. Frápolli and Esther Romero, CSLI Publications, 2003, pp. 1–31.

Gauker, Christopher. *Conditionals in Context*. Cambridge. MIT Press, 2005.

Gauker, Christopher. “The Circle of Deference Proves the Normativity of Semantics”. *Rivista di Estetica*, vol. 34, 2007, pp. 181–198.

Gauker, Christopher. *Words and Images: An Essay on the Origin of Ideas*. Oxford University Press, 2011.

Gauker, Christopher. “What Tipper is Ready For: A Semantics for Incomplete Predicates”. *Noûs*, vol. 46, 2012, pp. 61–85.

Gauker, Christopher. “The Illusion of Semantic Reference”. *On Reference*, edited by Andrea Bianchi, Oxford University Press, 2015, pp. 21–39.

Gauker, Christopher. “Inner Speech as the Internalization of Outer Speech”. *Inner Speech: New Voices*, edited by Peter Langland-Hassan and Agustín Vicente, Oxford University Press, 2018, pp. 53–77.

Gauker, Christopher. “Paradoxes of Truth-in-context-X”. *Philosophical Studies*, vol. 180, 2023, pp. 1467–1489.

Gauker, Christopher. “Amodal Completion: Mental Imagery or 3D Modeling? *Review of Philosophy and Psychology*, 2024. https://doi.org/10.1007/s13164-024-00740-1

Grice, H. P. “Meaning”, *Philosophical Review*, vol. 66, 1957, pp. 377–388.

Grice, H. P. “Utterer’s Meaning and Intentions”. *Philosophical Review*, vol. 78, 1969, pp. 147–177.

Harman, Gilbert. “Sellars’ Semantics”. *The Philosophical Review*, vol. 79, 1970, pp. 404–419.

Kaplan, David . “Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals”. *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein, Oxford University Press, 1989, pp. 481–564.

Murzi, Julien. “Classical Harmony and Separability”. *Erkenntnis*, vol. 85, 2020, pp. 391–415.

Peregrin, Jaroslav. *Inferentialism: Why Rules Matter*. Palgrave Macmillan, 2014.

Punčochář , Vít, and Christopher Gauker. “Indicative Conditionals in Objective Contexts”. *Theoria* (Sweden), vol. 86, 2020, pp. 651–687.

Putnam, Hilary. “The Meaning of ‘Meaning’”. *Minnesota Studies in the Philosophy of Science, Vol. 7: Language, Mind, and Knowledge*, edited by Keith Gunderson, University of Minnesota Press, 1975, pp. 131–193.

Rorty, Richard. *Philosophy and the Mirror of Nature*. Princeton University Press, 1979.

Sellars, Wilfrid. “Empiricism and the Philosophy of Mind”. *Science, Perception and Reality*. The Humanities Press, 1963, pp. 127–196. Originally published in: Feigl, Herbert and Michael Scriven, editors. *Minnesota Studies in the Philosophy of Science, Vol. 1: The Foundations of Science and the Concepts of Psychology and Psychoanalysis*. University of Minnesota Press, 1956, pp. 253–329. [EPM]

Sellars, Wilfrid. “Language as Thought and Communication”. *Philosophy and Phenomenological Research*, vol. 29, 1969, pp. 506–527. [LTC]

Sellars, Wilfrid. “Meaning as Functional Classification (a Perspective on the Relation of Syntax to Semantics)”. *Synthese*, vol. 27, 1974, pp. 414–437. [MFC]

Sellars, Wilfrid. *Science and Metaphysics*, Routledge & Kegan Paul Ltd, 1968. [SM]

Sellars, Wilfrid. “Some Reflections on Language Games”. *Science, Perception and Reality*. The Humanities Press, 1963, pp. 321–358. [SRLG]