JH
AP

Davidson on First-Person Authority and the Essential Sociality of Meaning

Ryo Tanaka
tanaka.r29@gmail.com
University of Tsukuba, Japan

ABSTRACT

As some authors have recently pointed out, the notion of *first-person authority* has an importantly social dimension: it concerns not only how one's mental self-ascriptions can be usually (or always) true, but also how one's *interlocutor* can presume such mental self-ascriptions to be true (Borgoni 2019, Winokur 2023). An adequate theory of first-person authority should account for both aspects of the phenomenon, but the traditional discussion has focused mainly on the former. In this paper, I will reexamine Davidson's account of first-person authority in this light and argue that it has interesting theoretical resources to address both aspects of first-person authority within a unified framework. In the literature, there is a consensus that (1) the key idea in Davidson's account is to explain first-person authority in terms of a speaker's *authoritative knowledge of the meanings of her own words*, and (2) Davidson explains the special status of a speaker's semantic knowledge by invoking his *non-communitarian* view of meaning-determination. I will argue that, to see a fuller picture of Davidson's view on first-person authority, these ideas need to be integrated with his discussion on triangulation, where he purports to show that meaning should be essentially *social* and *interpretable*. This will allow us to recognize how Davidson's theoretical system, as a whole, is indeed tailored to capturing the social dimension of first-person authority highlighted in the recent discussion.

1. Introduction: The Social Dimension of First-Person Authority

It is an essential part of our ordinary conception of self-knowledge that each person enjoys a distinctive epistemic security in ascribing occurrent mental states to the self. The distinctive feature of self-knowledge,

which Davidson ([1984](#)) calls *first-person authority*, is reflected in our pre-theoretical observation that one cannot be usually mistaken about what one is thinking about right now. Whenever one sincerely ascribes an occurrent thought to oneself by uttering, e.g., "I believe that Wagner died happily," there is normally a strong *presumption of truth*: if the speaker sincerely says she is entertaining a certain thought right now, then she usually (or always) has it. Correlatively, one's interlocutor is supposed to take the truth of the self-ascription for granted; and, under normal circumstances, it is even unclear how she might reasonably challenge it. Self-ascriptions of mental states are also characteristically *direct*: one can simply know one's own mind without relying on any evidence, and "How do you know?" is usually a non-sensical question to be raised against sincere mental self-ascriptions. One's ascriptions of mental states to *others*, on the other hand, do not share any of these features. They need to be based on evidence, they can often turn out to be false, and they can be easily challenged. Mental ascriptions to the self and to others thus seem to exhibit some distinctive kind of asymmetry. Many have attempted to explain it, including Davidson himself.[1,2]

As some authors recently point out, it is important to notice that the notion of first-person authority has an importantly *social* dimension (Borgoni 2019; Winokur 2023; Tanaka 2024; see also Schwengerer 2021 for a similar discussion). It contains not just the idea that one's sincere mental self-ascriptions are usually (or always) true despite the absence of evidence, but also the idea that one's *interlocutor* is supposed to treat such mental self-ascriptions as true and can thereby gain knowledge about the speaker's mind by simply accepting them. These two aspects of first-person authority are closely related but conceptually distinct. The following scenario provided by Borgoni ([2019](#)), for example, demonstrates this point vividly by showing how they could in principle come apart:

---

[1] In this paper, I will focus on one's knowledge of the contents of one's own thoughts and beliefs, following Davidson. In the interest of space, I will not examine to what extent the discussion can be generalized to cover other types of intentional mental states.

[2] Davidson's original formulation of the question is the following: "When a speaker avers that he has a belief, hope, desire or intention, there is a presumption that he is not mistaken, a presumption that does not attach to his ascriptions of similar mental states to others. Why should there be this asymmetry between attributions of attitudes to our present selves and attributions of the same attitudes to other selves?" (Davidson 1984, 3)

> Imagine yourself being a woman (or someone belonging to a gender, race, or other target group of prejudice). You tell your colleague, after having considered a candidate's application portfolio for an open position in your department, "I disagree with your proposal. I believe that this is the person we should hire." Your colleague, dismissing what you have just said and appealing to prejudicial ideas replies, "No, you don't believe it. Women just like to fuss with men's decisions" (Borgoni 2019, 296).[3]

In this case, although the woman's mental self-ascription, "I believe that this is the person we should hire," *is* true, her colleague *does not take it* as true (due to a sexist prejudice, in this case). As Borgoni puts it, first-person authority is "not granted unless our *peers* recognize it in our words" (2019, 298, emphasis mine).[4]

Apparently, however, the social dimension of first-person authority has not been the primary focus of the discussion in the literature. See, for example, the formulation of the notion of *first-person authority* found in Coliva (2016):

> Given C-conditions (including concepts' possession, cognitive well-functioning, alertness and attentiveness), if one judges to have a mental state M (save for dispositional ones or for the dispositional elements of some mental states), one will usually [or, always, in her stronger formulation] have it. (Coliva 2016, 65)

The formulation captures the idea that one's mental self-ascription is always or usually true despite the absence of evidence, but it does not explicitly introduce the perspective of the interlocutor. It does not mention the idea that one's interlocutor is supposed to take such mental

---

[3]See Schwengerer (2021, sec. 2), for a similar example.

[4]As Winokur (2023) points out, furthermore, it is ultimately inessential to draw on a case of epistemic injustice in the sense of Fricker (2007) to demonstrate this point (Winokur 2023, 7). To show this point, Winokur draws on the example given by Coliva (2019):

> I say, "I am healthy", while my spouse sees the number of pills I am taking and the disproportionate number of medical tests I undertake, and tells me, "you don't believe you are healthy, really. Look at the amount of pills you take and the numerous tests you undergo". (Coliva 2019, 344)

In this case, even if the agent's mental self-ascription is still true, the spouse can, apparently, raise reasonable doubts about its truth, given the inconsistency between the agent's self-ascribed belief and the actual behavior. This is enough to show that even if one's mental self-ascription is true, whether one's interlocutor is supposed to treat it as true remains as a separate question.

self-ascriptions to be true and can thereby gain knowledge about the speaker's mind. Both may usually go together, but they are nonetheless distinct and can come apart in some cases. The former concerns a speaker's ability to make true mental self-ascriptions, and as such, it captures the internal, or personal, dimension of first-person authority. The latter concerns an interlocutor's ability to gain knowledge about the speaker's mind by attaching the presumption of truth to her mental self-ascriptions, and as such, it captures the relational, or social, dimension of first-person authority.

A philosophical account of first-person authority should be able to incorporate both internal and relational aspects of the phenomenon. More specifically, one's account of first-person authority should be able to explain how a speaker's mental self-ascriptions can be usually (or always) true despite the absence of evidence, in such a way that it also explains, or *at least does not make unintelligible*, how one's interlocutor could acquire knowledge about the speaker's states of mind often enough by taking such mental self-ascriptions at face value (Tanaka 2024). With this in mind, it would be an interesting project to examine various accounts of first-person authority developed in the past few decades by asking whether it can meet this theoretical constraint that has started to attract some attention in the recent discussions.

In this paper, I will reexamine Davidson's account of first-person authority in light of this recent discussion and argue that it has interesting theoretical resources to address both the personal and the social aspects of first-person authority within a unified framework. Below, I will roughly explain the outline of the main discussion developed in the subsequent sections.

Among commentators on Davidson, there is a consensus that Davidson's official account of first-person authority invokes two key ideas. The first is the idea that first-person authority can be explained in terms of a speaker's *authoritative knowledge of the meanings of her own words*. The second is the idea that the special status of a speaker's semantic knowledge is to be explained in terms of a *non-communitarian externalist view of meaning-determination*. According to this view, roughly, the meanings of the words produced on the lips of an individual are generally determined by how that *individual herself* is disposed to apply them to objects in the environment. A speaker cannot be wrong about what her

words deployed in mental self-ascriptive utterances mean, because it is the speaker's own use that confers meaning to those words in the first place.

As I will argue, to see a fuller picture of Davidson's view on first-person authority, it needs to be integrated with his discussion on triangulation, where he purports to show that meaning should be essentially *social* and *interpretable*. My strategy for demonstrating this point will be to formulate an apparent puzzle for Davidson concerning the social dimension of first-person authority and explain how it can be resolved in his theoretical framework. Davidson's claim that meaning is determined in the non-communitarian way, when taken in isolation, may seem to suggest that there is no external social constraint on what a speaker can mean with her linguistic expressions. From this, one could be misled to think that Davidson's account will actually face the skeptical worry concerning knowledge about other minds: the contents of the beliefs that a speaker expresses by making self-ascriptive utterances could be *absolutely private* in the sense that no person other than the speaker herself can ever correctly articulate them. If so, one might think that even if Davidson succeeds in explaining how a speaker's mental self-ascriptions cannot be usually mistaken, he does so at the cost of *leaving it mysterious how a hearer can attach the presumption of truth to the speaker's mental self-ascriptions and thereby obtain knowledge about the speaker's mind*. In short, the worry is whether Davidson's account of first-person authority can successfully capture the social dimension of first-person authority. To resolve this problem, we will need to look at his discussion of triangulation. As we will see, Davidson's view is that a speaker could not mean anything by a word unless it is in principle *interpretable* by another person. This, by itself, may be an idea familiar to the readers of Davidson. However, Davidson does not elaborate on the social aspect of his account of meaning in the papers where he discusses the issue of first-person authority, and commentators have not examined exactly how these two issues are connected in Davidson's philosophical system, taken as a whole. The aim of this paper is to fill this gap in the literature and show how Davidson's view, when examined holistically, is indeed tailored to capturing the social dimension of first-person authority.

The discussion proceeds as follows. In Section 2, I will explain the two key ideas that constitute Davidson's official account of first-person authority more carefully. In Section 3, I will formulate the skeptical worry concerning knowledge of other minds that might seem to follow from Davidson's non-communitarian view of meaning-determination. In Section 4, drawing on his discussion of triangulation, I will explain how Davidson's view of meaning-determination is crucially social in its own way and argue that it has an important implication on the issue of interpretability. In Section 5, I will connect Davidson's view on the social nature of meaning with his account of first-person authority to demonstrate how this allows Davidson to dissolve the skeptical worry and thereby incorporate the social dimension of first-person authority. Section 6 is the conclusion.

2. Davidson's Account of First-Person Authority

2.1. First-person authority explained by speaker's authoritative knowledge of meaning

There is a consensus among commentators that Davidson's basic explanatory strategy as developed in Davidson (1984) and (1987) ("First-person Authority" and "Knowing One's Own Mind") is to show how we can understand the phenomenon of first-person authority concerning mental self-ascriptions in terms of the authoritative *knowledge of meaning* that a speaker possesses with regard to her self-ascriptive utterances, such as "I believe that Wagner died happily".[5] This is the first key component in Davidson's account of first-person authority. Here is how he sets up the discussion: to know what belief a speaker expressed by a certain utterance, in general, one should know the meaning of the sentence uttered, as well as whether the speaker takes the sentence to be true. This is how Davidson himself puts it: "if you or I or anyone knows that I hold this sentence true on this occasion of utterance, and she knows what I meant by this sentence on this occasion of utterance, then

---

[5]The role of knowledge of meaning in Davidson's account of first-person authority is discussed by many authors including Bar-On (2004, 173–86), Beisecker (2003), Child (2007); Child (2013); Child (2017), Hacker (1997), Higginbotham (1998), Jacobsen (2009), Ludwig (1994), Macdonald (1995), Picardi (1993), Smith (1998, 417), Thöle (1993), Winokur (2021), Wright (2001), Yalowitz (1999).

she knows what I believe—what belief I expressed" (Davidson 1984, 11). Davidson, for the sake of discussion, simply assumes that both the speaker and the hearer know that the speaker takes the sentence to be true. Then, he proposes to explore whether first-person authority could be explained by the idea that the speaker can know the meanings of one's own words in some distinctively secure way that is not available to the hearer.

Davidson proposes that first-person authority derives "from the fact that the assumption that I know what I mean necessarily gives me, but not you, knowledge of what belief I expressed by my utterance" (1984, 12). More precisely, Davidson points out that we should look at a distinctive *asymmetry* that holds between a speaker's knowledge of the meaning of her own self-ascriptive utterance and a hearer's knowledge about it: there is "a presumption that speakers, but not their interpreters, are not wrong about what their words mean" (1984, 12). Davidson notes that this is not to say that speakers are *never* wrong about the meanings of their words (1984, 12). Rather, the relevant asymmetry is that speakers cannot be *regularly and systematically* mistaken about the meanings of their words, whereas that is perfectly possible for their interlocutors. This, on Davidson's view, is the source of first-person authority: when a speaker ascribes a belief to herself by making a self-ascriptive utterance, "I believe that Wagner died happily," the speaker knows the meaning of "Wagner died happily" as uttered by herself in some distinctively secure way that is not available to the hearer. So, Davidson claims, the speaker knows what belief is ascribed to herself by the utterance in question, in some distinctively secure way that is not available to the hearer. This is how he summarizes his proposal: "[there is] a presumption. . .that the speaker usually knows what he means. So there is a presumption that if he knows that he holds a sentence true, he knows what he believes" (1984, 14).

To complete his account, however, Davidson needs to explain the *source* of the presumption that a speaker is not regularly mistaken about the meanings of her own words. Without some such explanation, it would seem that Davidson's account simply replaces one philosophical problem (i.e., first-person authority concerning belief self-attributions) with another (i.e., a speaker's authoritative knowledge of her own meaning). In Davidson (1984), he states that the asymmetry in question

is reflected in the idea that a speaker can know the meaning of her own utterance *directly*, whereas the hearer must rely on evidence.[6] This, however, seems to provide a mere description of the relevant asymmetry, rather than an explanation.[7] Commentators point out that Davidson's explanation of a speaker's authoritative knowledge of meaning—the second key component in his overall account of first-person authority—can be found in his later writings on this topic.

2.2. Speaker's authoritative semantic knowledge explained by non-communitarian semantic externalism

Here is a key passage from Davidson (1987):

> The explanation [of first person authority] comes with the realization that what a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the words to be applicable. . .An interpreter of another's words. . .must depend on scattered information, fortunate training, and imaginative surmise in coming to understand the other. The agent herself however, is not in a position to wonder whether she is generally using her own words to apply to the right objects and events, since whatever she regularly does apply them to gives her words the meaning they have and her thoughts the contents they have. . .To put the matter another way, nothing could count as someone regularly misapplying her own word. (Davidson 1987, 37–38)

The passage contains two important claims that need to be unpacked. Let me explain each in turn.

---

[6]Here is the relevant passage:

A hearer interprets (normally without thought or pause) *on the basis of many clues*. . .The speaker, although he must bear many of these things in mind when he speaks, since it is up to him to try to be understood, cannot wonder whether he generally means what he says. (Davidson 1984, 12, emphasis mine)

See also p. 13 for a similar passage.

[7]Elsewhere, Davidson claims that the presumption that each speaker knows her own meaning constitutes a transcendental condition for the possibility of interpretation. See, for example, the following passage: "unless there is a presumption that the speaker knows what she means, i.e., is getting her own language right, there would be nothing for an interpreter to interpret" (Davidson 1987, 38). It is a matter of dispute whether this should be taken as Davidson's official account of first-person authority, instead of the one that is grounded in his account of meaning-determination (see Child 2007, sec. 3.i; Hossein Khani 2021). I will return to this issue in Section 5.2.

As Child (2007) helpfully observes, for Davidson, "the fundamental justification for the presumption that speakers generally know what their words mean comes from considerations about meaning-determination" (Child 2007, 165). Importantly, what is at work here is the view that we might call *non-communitarian externalism*, according to which the meanings of words uttered on the lips of a speaker are determined generally by what objects and events that the speaker regularly applies the words to. This is a form of *externalism*, because the meanings of words in a speaker's language are determined by causal interactions between the speaker and objects/events external to her.[8] Furthermore, the view is *non-communitarian* in that the meanings of words in a certain speaker's language are determined generally by that speaker's patterns of language use. That is, Davidson's externalism is crucially different from the Burgean *communitarian* externalist view, according to which the meanings of one's words are determined in part by how *other speakers* in the same linguistic community tend to use them (Burge 1979).

So, the first key idea presented in the above-cited passage is this: a speaker cannot make regular errors in using her own words, *because* their meanings are determined generally by the speaker's overall patterns of use in the first place. To use an analogy by Child: a person cannot regularly act out of character because the person's character is determined by how she generally acts; similarly, a speaker cannot regularly misuse her words because their meanings are determined by how the speaker generally uses them (Child 2013, 539). The hearer's usage of the same-sounding words, on the other hand, does not similarly contribute to the meaning-determination of the speaker's words. Davidson thinks that this asymmetry in meaning-determination explains the distinctive security of a speaker's knowledge of the meanings of one's own words.

Indeed, Davidson contends that those who adopt the Burgean communitarian externalist view would have a serious problem in explaining first-person authority. As is well-known, the Burgean view allows for the possibility that a competent speaker can be radically mistaken about the

---

[8] As Jacobsen puts it, Davidson's view is that "the words uttered by a speaker are constituted as meaning what they do by their relations to objects and events in the speaker's external environment" (Jacobsen 2009, 255). See e.g., Beisecker (2003), Macdonald (1995), Child (2007), Child (2013), Child (2017), Winokur (2021), and Yalowitz (1999) for similar discussions of the role of Davidson's view of meaning-determination in his account of first-person authority.

meaning of a term in her vocabulary and regularly use it in a confused way (e.g., "I have arthritis in my thigh," where "arthritis" means *joint inflammation*). Davidson explicitly claims that this would undermine a speaker's authoritative knowledge of her own meaning (Davidson 1987, 26–27).[9] Reflecting this, in a later article, Davidson notes that he endorses "a fairly extreme form of individualism about meaning," which is "not accepted by those who think languages are defined by shared practices" (Davidson 1993a, 250). On Davidson's view of meaning-determination, the same-sounding words produced on the lips of different speakers may, and often do, have different meanings in their idiolects. Of course, they might sometimes converge with one another to some extent. But Davidson's point is that they need not.[10]

Now, let us move on to the second key point. Construed in this way, there seems to be an immediate objection to Davidson's explanation: even if a speaker's words in her language are determined generally by the speaker's own use, it does not show that the speaker is thereby in a privileged position to *know what their meanings are*. To use Child's analogy again: even if a person character is determined by how she generally acts, she may not necessarily know what her character is—it requires observing her actual behaviors in various circumstances (Child 2013, 539). One way to respond to this objection on behalf of Davidson is to think that he endorses what we may call the *deflationary* conception of knowledge of meaning. On the deflationary conception, a speaker's knowledge of the meanings of her own words can be identified with the speaker's *ability* (*knowing-how*) to use them correctly, instead of the speaker's possession of some propositional knowledge *about* their

---

[9]Here is the passage:

...I am convinced that if what we mean and think is determined by the linguistic habits of those around us in the way Burge believes they are, then first person authority is very seriously compromised...I agree that what I mean and think is not 'fixed' (exclusively) by what goes on in me, so what I must reject is Burge's account of how social and other external factors control the contents of a person's mind. (Davidson 1987, 26–27)

See also Davidson (1997, 199).

[10]Davidson's claim may not be surprising given Davidson's general view of language, which is epitomized in his famous claim that there is no such thing as language if by "language" one means a system of meaningful symbols that is *necessarily* shared by all speakers who can successfully communicate with one another (Davidson 1986). I will return to this issue in Section 4.

meanings.[11] This is to think that Davidson understands a speaker's authoritative knowledge of the meanings of her own words in terms of the speaker's *ability to use them in a way that is immune to regular errors*. As he puts it in the passage cited above, "nothing could count as someone regularly misapplying her own word".[12] Arguably, this kind of linguistic/semantic knowledge-how that constitutes a speaker's linguistic competence need not consist in the propositional knowledge *that one's words mean such and such*.[13,14] If this is the conception of semantic knowledge that Davidson needs, then the objection can be avoided. (To clarify, some commentators such as Child argue that we could also find a richer, non-deflationary conception of semantic knowledge in Davidson. Although the plausibility of the discussion below will not hang on this issue, see footnote for discussion).[15]

---

[11]Here are some other passages that support the attribution of the deflationary conception of semantic knowledge to Davidson:

> The word 'know' here is not essential in making the point; it would do as well to say that the speaker is not misusing his own words, or that he means, in a literal sense, what he says. (Davidson 1993b, 212)
>
> [F]or the sake of making this point, one might well decline to use the word 'knowledge' for what accounts for our correct use of words to express our thoughts. (Davidson 1993a, 249)

Jacobsen (2009, 264 n5) also cites these passages.

[12]Commentators thus point out that the distinctive epistemic status of a speaker's knowledge of meaning is spelled out in terms of the speaker's "immunity to regular errors of misapplication" (Jacobsen 2009, 255)—in other words, this is to think that "knowing the meaning of a word just consists in the ability to use it correctly" (Child 2013, 539).

[13]Bar-On helpfully summarizes the overall structure of Davidson's account as follows: "we can understand the notable first-person/other-person asymmetries in propositional attitude ascriptions, if we reduce knowledge of the contents of one's propositional attitudes to semantic knowledge, and reduce semantic knowledge, in turn, to competent language use" (Bar-On 2004, 177).

[14]I acknowledge that one could challenge the Rylean distinction between knowledge-that and knowledge-how presupposed in the discussion (see Stanley and Willlamson (2001)). Providing a defense of the distinction is beyond the scope of the current paper, and I shall merely point out that the distinction is still widely used and the commentators on Davidson do not take issue with it.

[15]As I explain in the next paragraph, Davidson does say that a competent speaker could provide a disquotational articulation of the meanings of her own words (Davidson 1984, 13). This suggests that Davidson thinks that a competent speaker is equipped with genuine propositional knowledge about meaning, although the content may be purely disquotational. Child examines how, in Davidson's framework, a competent speaker can come to acquire disquotational propositional knowledge of meaning, based on her basic semantic competence plus minimally required meta-linguistic abilities (Child 2013, 540).

To see how this allows Davidson to bypass the objection, it should be helpful to notice how it explains the required kind of *directness* of a speaker's authoritative knowledge of meaning. A speaker's authoritative knowledge of meaning (construed in a deflationary way) need not be supported by evidence, Davidson can say, because the knowledge involves no proposition to be supported in the first place. On the deflationary reading, roughly, *being* a competent user of a language suffices for *knowing* one's own meanings. Just as one can know how to ride a bicycle without possessing any articulate propositional knowledge that necessary maneuvers are such and such, one can know how to correctly use words without possessing any propositional knowledge that their meanings are such and such. Indeed, Davidson seems to think that a fully competent speaker need not be able to articulate the meanings of her own words, other than in the purely disquotational form: "[t]he speaker, after bending whatever knowledge and craft he can to the task of saying what his words mean, cannot improve on the following sort of statement: 'My utterance of 'Wagner died happy' is true if and only if Wagner died happy'" (Davidson 1984, 13).

Let me summarize the overall structure of Davidson's official account of first-person authority. The explanandum is the asymmetry that exists between a speaker and a hearer: by uttering a self-ascriptive utterance, the speaker can ascribe a belief to herself in the distinctively secure way that is not available to the hearer. Davidson thinks that the asymmetry can be explained in terms of the speaker's knowing the meanings of her own words in the distinctively secure way that is not available to the hearer. On Davidson's view, the asymmetry in semantic knowledge can be explained by his non-communitarian externalist view of meaning-determination: the speaker cannot regularly misuse her own words because it is the speaker's regular use that determines their meanings, whereas a hearer does not similarly contribute to the determination of their meanings. Lastly, in order for Davidson's account to work, apparently, he needs to adopt a deflationary conception of speaker's knowledge of meaning.

In the rest of the paper, I will assume that this provides a by and large accurate reconstruction of Davidson's official account of first-

person authority.[16] Below, I will argue that it is crucial to connect Davidson's account of first-person authority with his fuller account of meaning-determination that draws on the triangulation argument, because failing to do so can mislead us to have the impression that due to his commitment to "a fairly extreme form of individualism about meaning," Davidson in fact fails to incorporate the social dimension of first-person authority in his framework. In the next section, I will explain this point more carefully.

3. Skeptical Worry for Davidson?: Sharing Beliefs through Self-ascriptions

For the sake of discussion, I shall raise an objection to Davidson's account of first-person authority along the following lines. His non-communitarian externalism, one might say, will prevent him from accommodating the social dimension of first-person authority we observed in Section 1. This is because the view in tension with what we may call the *ordinary view* concerning a *hearer's* acceptance of a speaker's self-ascriptive utterance. To introduce the ordinary view, consider this simple question: for a speaker, what is the *point* (or, *purpose*) of making a self-ascriptive utterance, after all? The point, at least in part, consists in allowing a *hearer* to acquire knowledge about the speaker's current mental state through her comprehension of the speaker's self-ascriptive utterance. Linguistically ascribing a mental state to the self is a way of *sharing*, as it were, the self-ascribed mental state to a linguistically competent hearer.[17] As observed in Section 1, a certain methodological

---

[16]In this paper, I also assume that Davidson's official view of meaning-determination does not rely on the Gricean idea that what determines the meanings of one's words is ultimately one's communicative intentions. Admittedly, this is debatable—even in Davidson (1984) and (1987), there are passages where he invokes the notion of intention (see Davidson 1984, 14; 1987, 28). However, if Davidson's view were Gricean, it would make his account of first-person authority circular: since intentions are propositional attitudes like beliefs, an account of one's knowledge of one's own beliefs, apparently, should not presuppose that one has privileged access to one's own intentions. For this reason, I shall assume that the notion of communicative intention does not (should not) play any direct explanatory role in Davidson's official account. Child, for example, adopts a similar stance toward this issue (Child 2007, 164).

[17]Bar-On (2004, 264) articulates this kind of idea nicely, by arguing that mental self-ascriptive utterances typically exhibit the feature that she calls "transparency-to-the-

constraint that any adequate account of first-person authority should aim to meet: one's account should explain first-person authority concerning a speaker's mental self-ascriptions, in such a way that it also explains, or at least does not make unintelligible, how a *hearer* can attach the presumption of truth to the speaker's mental self-ascriptions and thereby gain knowledge about the speaker's mind by accepting them at face value.[18]

Those who find the ordinary view plausible would describe what is typically involved in a speaker's self-ascriptive utterance and a hearer's acquisition of knowledge about the speaker's mind in the following way. Suppose that a speaker, Ann, ascribes a belief to herself by uttering, "I believe that Wagner died happily." Suppose also that Ben, her interlocutor, hears every single word of her utterance correctly, and all of the words Ann uttered sound perfectly familiar and intelligible to his ears. Ben recognizes Ann's first-person authority, and, assuming that Ann is sincere, Ben proceeds to take Ann's words for granted. As a result, this should put Ben in a position to legitimately say, "You believe that Wagner died happily," and acquire knowledge that Ann believes that Wagner died happily. There should be nothing puzzling here. After all, the *purpose* of Ann's making the mental self-ascription is presumably to share her mental state with Ben, as highlighted in the recent discussion on the social dimension of first-person authority.[19]

Here, one may question whether Davidson can adequately explain the legitimacy of this ordinary practice in his framework. The question can be motivated by recognizing the fact that, in ascribing a belief to Ann, Ben should use *his own words* to specify the content of Ann's belief. Apparently, if we adopt Davidson's view of meaning-determination, Ben is not necessarily entitled to the presumption that he and Ann mean

---

subject's-state": by making a mental self-ascription, a speaker is allowing others to *see through to* the state ascribed to oneself.

[18] See Tanaka (2024) for an elaboration of this point.

[19] Also, recall that Davidson's initial characterization of first-person authority draws on the observation that "there is a presumption of truth" concerning one's mental self-ascriptions (Davidson 1984, 3), which naturally leads to the idea that when a speaker makes a sincere self-ascriptive utterance, a hearer should be allowed to *presume that the speaker's self-ascription is true* and can acquire a justified belief by relying on it. This is to think that the phenomenon of first-person authority is reflected in the practice where a speaker's sincere mental self-ascription gives a hearer a default entitlement to take it at face value and form a justified belief about the speaker's mental state.

the same things by the same-sounding words, after all. Since Davidson rejects the Burgean communitarian externalism, Ann and Ben may be using the same-sounding words with different meanings in their own idiolects. If this is so, to acquire knowledge about Ann's mental state, Ben would have to first confirm whether the meanings he attaches to the relevant words (i.e., "Wagner," "died," "happily") match the meanings Ann attaches to the same-sounding words in her language.[20]

So far, this by itself does not pose any serious problem yet. Occasional misunderstandings may be simply common. Indeed, Davidson's point is precisely that there is such a speaker-hearer asymmetry: while a speaker cannot be generally mistaken about the meanings of her own words, a hearer can.[21] However, by pressing this line of thought further, we will face an important question. In Davidson's framework, is it guaranteed that Ben could *in principle* correctly specify the meanings of the relevant words in Ann's language, and thereby correctly articulate the content of the belief Ann ascribes to herself? Or, is it possible that, no matter what evidence Ben manages to collect concerning Ann's linguistic behaviors, Ben could never be able to correctly articulate the content of Ann's belief, using his own words? The question concerns whether Davidson's non-communitarian view of meaning-determination may prevent him from providing a plausible explanation of how a speaker can come to share her belief with a hearer via language—in other words, how a speaker can make available the content of her current belief to a speaker by making a self-ascriptive utterance.

The question could turn into a skeptical worry, if one thinks that in rejecting the Burgean communitarian externalism, Davidson subscribes to the view that a person's having a history of applying symbols to objects and events in the environment in some regular ways is *sufficient* for the symbols to have determinate meanings. This, in effect, would imply that a speaker's idiolect could emerge as a solitary language: a language that one develops only for herself from scratch without relying on any social interaction. If this is Davidson's view, it seems that he

---

[20]Goldberg (2004) offers a similar observation in his critique of Davidson's discussion of radical interpretation.

[21]Beisecker claims that the speaker-hearer asymmetry that Davidson posits poses a problem itself and argues that we should reject it (Beisecker 2003, 92). For the purposes of the current paper, we need not examine the plausibility of Beisecker's claim, which amounts to the wholesale rejection of Davidson's account of first-person authority.

leaves it unexplained how a hearer can come to acquire knowledge about the speaker's current state of mind by taking the speaker's self-ascriptive utterance as true. For, if expressions used in the speaker's self-ascriptive utterances belong to her solitary language, it is unclear how others can come to use them with the meanings they have in the speaker's language. In the worst case scenario, it might even seem that, for a hearer to understand the solitary language developed by the speaker, she would need to become the speaker herself—that is, for two people to use the same sounding words with the same meanings, they would have to share the exact same history of applying those words to objects and events in the environment. If so, it would become realistically impossible for any two speakers to share meaning. This, if true, would render entirely ungrounded what I described as the ordinary view concerning a hearer's acceptance of a speaker's self-ascriptive utterance.

If Davidson's account of first-person authority is indeed susceptible to this kind of skeptical worry, this would be a serious flaw: even if it explains the personal dimension of first-person authority, it does so at the cost of making its social dimension completely unintelligible. Of course, as I will explain, Davidson has a plausible answer to this worry. This becomes clear once we recognize that Davidson in fact *does not* think that a person's having a history of applying symbols to objects and events in the environment in some regular ways is sufficient for the symbols to have determinate meanings. However, crucially, if we focus exclusively on his remarks on the issue of meaning-determination offered in his discussion of first-person authority, we can be misled to think otherwise. In the key passage I cited in the previous section, Davidson claims that "what a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the words to be applicable," and "whatever she regularly does apply them to gives her words the meaning they have and her thoughts the contents they have" (Davidson 1987, 37). We have also seen that he describes his position as "a fairly extreme form of individualism about meaning" (Davidson 1993a, 250). From these remarks, one might take Davidson to be committed to the idea that a person's having a history of applying symbols to objects and events in the environment in some regular ways is sufficient for the symbols to have determinate meanings.

Furthermore, some of the expository remarks on Davidson's account of first-person authority made by commentators might also help generate the impression that Davidson's view of meaning-determination is not only non-communitarian but thoroughly non-social. See, for example, the following passages from Jacobsen (2009) and Child (2013):

> . . .if, as Davidson believes, immunity to error can be assigned to individual speakers, then each solitary speaker's uses of her words must be assumed to confer meaning on them. A speaker's immunity to regular verbal error is then an expression of the idea that only her own uses of words can create the word-world patterns needed to constitute her utterances of those words as meaning what they do. (Jacobsen 2009, 256)

> The main substance of the explanation comes from the idea that the meanings of a speaker's words are determined by her own use of those words. That is what underpins the presumption that a speaker usually knows what her words mean. (Child 2013, 534)

These passages can help generate the impression that Davidson is talking about a solitary language in his discussion of meaning-determination. At least, in their expositions of Davidson's account of first-person authority, neither Jacobsen nor Child connect it with Davidson's fuller account of meaning-determination developed elsewhere. Yalowitz (1999), even more explicitly, claims that no explanatory role needs to be given to social interactions in Davidson's account of first-person authority and claims that it must "be applicable wholesale to the solitary language user" (Yalowitz 1999, 111).[22]

To clarify, these commentators, with the exception of Yalowitz, do not explicitly attribute the thoroughly non-social account of meaning-determination to Davidson. Nonetheless, it also has to be noted that they do not attempt to connect Davidson's account of first-person authority with his discussion on the sociality of meaning either. As a result, I think their expositions can help generate the false impression that Davidson's account of first-person authority is subject to the skeptical worry concerning knowledge about other minds that I formulated above and thereby fails to incorporate the social dimension of first-person authority.

---

[22]To clarify, Yalowitz's aim in his paper is to argue against Davidson's account of meaning-determination. Since the aim of the current paper is to provide a reconstruction of Davidson's view, we need not look at Yalowitz's criticism of Davidson's view in detail.

In the rest of this paper, I will explain how Davidson's account of first-person authority is not in fact susceptible to the skeptical worry sketched above. To see this, we need to understand how Davidson's fuller account of meaning-determination incorporates a social element despite being non-communitarian. I turn to this task in the next section.

4. Davidson's Social Externalism: Triangulation Argument

My aim in this section is to provide a fuller reconstruction of Davidson's account of meaning-determination only to the extent that it is necessary to respond to the apparent problem I raised for Davidson in the previous section. The idea that meaning should be essentially social constitutes, in fact, an important strand in Davidson's philosophy. In his earlier writings, he develops it with the notion of radical interpretation; in his later writings, he prefers to do it with the notion of triangulation. I do not intend to cover every aspect of Davidson's view on meaning, nor do I aim to offer a completely novel interpretation of Davidson on this issue. Rather, my aim is to demonstrate how his view on the essential sociality of meaning and his account of first-person authority are connected in important places. This will allow us to see how his view, examined holistically, is in fact tailored to understanding the social dimension of first-person authority.

From the discussion in Section 2, we already know that Davidson's account of meaning-determination is *externalist*: "what a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the words to be applicable" (Davidson 1987, 37). Elsewhere, he calls this idea "*perceptual externalism*" (Davidson 1990, 198) and notes that he is strongly committed to it (see also Davidson 2001a, 2).[23] Also, as I explained, Davidson's view is

---

[23]Davidson says:

. . .for some thirty years [I have] been insisting that the contents of our earliest learned and most basic sentences ('Mama', 'Doggie', 'Red', 'Fire', 'Gavagai') must be determined by what it is in the world that causes us to hold them true. It is here, I have long claimed, that the ties between language and the world are established and that central constraints on meaning are fixed; and given the close connections between thought and language, analogous remarks go for the contents of the attitudes. (Davidson 1990, 200)

non-communitarian: the meanings of words in a speaker's idiolect need not be affected by other speakers' usage. Again, the question is whether Davidson thinks the existence of such causal connections between a person's words and objects and events in the environment is *sufficient* for meaning-determination.

As Davidson himself writes in the introduction of *Subjective, Objective, Intersubjective,* another key component in Davidson's externalism is the idea that "language is necessarily social," in the sense that "to have thoughts, and so to mean anything in speaking, it is necessary to understand, and be understood by, a second person" (Davidson 2001b, xv). In "Second Person," Davidson puts the idea in this way: "[T]here would be no saying what a speaker was talking or thinking about, no basis for claiming he could locate objects in an objective space and time, without interaction with a second person" (Davidson 1992, 121), and this shows that "one's first language cannot be a private language, that is, a language understood by only one creature" (Davidson 1992, 121).[24] In other words, Davidson's view is that establishing causal connections between a person's words and objects and events in the environment by itself is *not* sufficient for meaning-determination. Or, more accurately, it is not sufficient for meaning-*constitution*, because Davidson thinks that, at least in basic cases, there would be nothing that a speaker could talk or think about in the first place unless the speaker interacts with, and is understood by, other people. So, although Davidson rejects the Burgean communitarian externalism, his account of meaning-determination is social in its own way.[25] Why does Davidson think that meaning-constitution/determination requires interpersonal

---

[24]Here is another passage where Davidson puts the same point:

The problem is not, I should stress, one of verifying what objects or events a creature is responding to; the point is that without a second creature responding to the first, there can be no answer to the question. (Davidson 1992, 119)

[25]This is explicit in the following remark on Burge:

I have rejected social externalism as championed by Burge. Do I therefore think social factors play no role in the externalism of the mental? Not at all. But I would introduce the social factor in a way that connects it directly with perceptual externalism, thus locating the role of society within the causal nexus that includes the interplay between persons and the rest of nature. (Davidson 1990, 200)

interactions? To answer this question, we need to look at his discussion of triangulation. Below, I will mostly draw on Davidson (2001a), "Externalisms," where we can hope to find the most elaborate and accessible version of his triangulation argument in comparison to his earlier writings.[26]

Again, Davidson endorses the idea that the meanings of a speaker's words are determined by the typical causes of their tokenings. However, Davidson now claims that this does not give us a complete account of meaning-determination, because such "typical causes" of the tokenings of words can be grouped in a numerous different ways. Here is how Davidson puts the problem:

> The problem of relevant similarity comes up also in connection with the idea that the content of a perceptual belief depends on its common or usual cause. Even to approach the question, we must have some idea which causes are being gathered together. . . any set of causes whatsoever will have endless properties in common. . . (Davidson 2001a, 5)

Perhaps the best way to understand this problem is to formulate it via the Wittgensteinian observation that in order for a word to have a determinate meaning, some objective condition for its correct application should be established (Davidson 2001a, 3–4). If, for example, "cow" means *cow* on a speaker's lips, this makes the speaker's application of "cow" to an object correct (at least in basic cases) if and only if the object is *in fact* a cow; and, this simultaneously makes the speaker's application of "cow" to a fake cow incorrect, *however similar it might be to real cows*. Fake cows might be indeed very similar to cows, but this similarity is not "relevant," in the sense that they should not be categorized as "cow" on the basis of this similarity. As Davidson says, "[f]ake cows are in one way relevantly similar to real cows—that's why we make mistakes. But in another way, fake cows aren't at all like cows—they don't fall under the concept cow" (2001a, 3–4) Now, the question is this: what establishes the fact that the speaker's application of "cow" to a fake cow is incorrect, even if it looks highly similar to real cows for the speaker and she even has a disposition to apply "cow" in such a case? What is the "relevant similarity" that a group of objects, to which "cow" (as the speaker uses

---

[26]Examining connections between his triangulation argument and his earlier discussion of radical interpretation is also an important task, which I cannot undertake in this paper due to the limit of space (see Verheggen 2017 for a useful discussion on this matter).

it) is correctly applicable, must exhibit? And, more importantly, how are we to answer this question? The task, in effect, is to explain what it takes for a person to possess any *concept* at all (2001a, 4).

In outline, Davidson's proposal is that an adequate solution to the problem of "relevant similarity" requires *triangulation*: two people interacting with one another and with the world. In other words, he thinks that a solitary person who does not interact with another person cannot answer this problem.[27] Davidson notes that he is convinced by Wittgenstein (or, Kripke's (1982) discussion inspired by Wittgenstein) that "only social interaction brings with it the space in which the concepts of error, and so of meaning and of thought, can be given application" (2001a, 4).[28] To secure the possibility of genuine error/misapplication, Davidson argues, whether a speaker's reaction to an object is "relevantly similar" to her past reactions should not be a matter that the speaker alone can decide from the solipsist perspective. If this matter were entirely up to the speaker's subjective judgement, the distinction between what *seems* to be an instance of correct application for the speaker and what *is* in fact an instance of correct application would collapse; and, the possibility of genuine error/misapplication would be lost.[29] On Davidson's view, the distinction in question becomes available to a person only if she interacts with another person in the environment.

More specifically, Davidson explains how triangulation allows a person to recognize the possibility of genuine error/misapplication in the following way (2001a, 6). Imagine that there are two people, A and B, who can observe one another's reactions to objects in the environment. Suppose that A regularly responds to cows by making the sound "cow"; now, we have another person B, who can observe that there are such correlations between "cow" uttered by A and a certain group of objects,

---

[27]There is a complication here: should we really ask (as Davidson does) how a speaker *herself* can solve the problem of relevant similarity? Isn't it enough that *we* (theorists observing the speaker) can solve the problem for her? Davidson, however, insists that it is crucial that each speaker herself is in a position to recognize the distinction between correct and incorrect applications of her own words (Davidson 2001b, 6–8). This is continuous with his insistence that possessing the concept of objectivity should be a prerequisite for possessing any concept at all (see e.g., Davidson 1997, 8). See part I of Myers and Verheggen (2016) for an extensive discussion of this matter.

[28]See also Davidson (1991, 212–13); Davidson (1992, 119); Davidson (1997, 129).

[29]See Verheggen (2017, 153) for a fuller discussion of the current point. See also Sultanescu and Verheggen (2019).

which do look similar with one another to B herself (2001a, 6). Now, B is in a position to "notice occasions on which the responses of the other differs from the other's usual responses" (2001a, 6). That is, B can recognize when A applies "cow" to an object that, *from B's perspective*, A should not categorize as "cow," given A's past usage of the term. In this situation, "A's responses are not taken by B to be the same as A's earlier responses, while B's responses are the same as B's earlier responses" (2001a, 6).[30] From her own perspective, B can judge A's particular application of "cow" to an object in the current situation to be incorrect, in the sense that it is not relevantly similar to A's past reactions. Davidson's main idea is that, to secure the possibility of genuine error, it is crucial that there be such a judgement that could *diverge* from A's judgement about the correctness of her own application of the term. That is, even if A, *from her own perspective*, thinks that her application is in accordance with her past usage and thus is correct, it should be possible that there could be a genuine challenge to such a judgement. Such a challenge, on Davidson's view, can only be issued by another person, B, who has been observing A's past applications of "cow" to objects in the environment.

Davidson argues that this triangle allows A (and B, because A is also in a parallel position to do all of the above for B) to recognize the possibility of genuine error. This would "make the difference between the cases where the creatures act in unison, and the cases where they do not, available to the creatures themselves" (Davidson 2001a, 6). That is, A (as well as B) is now in a position to recognize that even if *she* thinks that the object in front of her should be categorized as "cow," " it might be mistaken. Although it is her past usage of "cow" that constitutes what it means for her, it is not entirely up to her to decide whether *her particular application of "cow" in a given situation* is correct or incorrect—that is, whether it is "relevantly similar" to her past usage. Davidson takes this triangulation argument to show that establishing the possibility of genuine error—that is, distinguishing "between a mere disposition and

---

[30] As Bernecker notes, this is, roughly, to think that "the similarity-standard for responses to a nearest mutual typical cause is that the interpreter judges the responses to be similar" (Bernecker 2013, 451). One important complication, however, is that B could also make mistakes. So, it is not that A should always correct her usage when B issues a challenge to A. Rather, Davidson's point is that the possibility of divergence on this matter between two creatures creates the possibility of genuine error.

rule following" (2001a, 7)—requires two people interacting with one another in the shared environment. As he puts it, "an interconnected triangle such as this. . .constitutes a necessary condition for the existence of conceptualization, thought, and language" (2001a, 7).

For the purposes of this paper, it is crucial to see that, if the triangulation argument is successful, it imposes a social constraint on what a speaker can mean by her own words: they need to be *interpretable* (and in basic cases, have been actually interpreted) by another person. On my reconstruction of Davidson's triangulation argument, the key idea is that the question whether A's current application of a word is consistent with her past usage can be genuinely raised only if there is another person, B, who can challenge A's own judgement on this matter. This, in turn, requires that A and B could at least agree on what A's word means—e.g., whether "cow" should be applied to this particular object of their joint attention, given A's past usage. In other words, although the meaning that "cow" has for A is constituted by A's own use of the symbol (and need *not* be affected by B's use of the same symbol), there is a crucial social constraint that another person B should be able to grasp the condition for correct application that "cow" has for A (i.e., what "rule" A is following in applying "cow"), based on her observations of A's regular applications of "cow" to objects in the environment. This is why "what a speaker means by what he says, and hence the thoughts that can be expressed in language, are not accidentally connected with what a competent interpreter can make of them (2001a, 11).[31]

To clarify, in this picture, there is no requirement that A and B should use their words with the same meanings. For example, there is nothing that prevents B from using the same sound, "cow," in whatever way she likes, *insofar as her own use is also interpretable by A*. Of course, it would certainly facilitate communication if A and B use "cow" with the same meaning. But they need not. Thus, the triangulation argument is in line with Davidson's rejection of the Burgean communitarian externalist view.

To summarize: as we have seen in Section 2, Davidson endorses the idea that the meaning of a speaker's word is determined by her past applications of the word to objects and events in the environments

---

[31]Thus, as Verheggen observes, the key feature in Davidson's account of meaning-determination is the idea that "the physical side of externalism can be secured only through the social side" (Verheggen 2017, 145).

(i.e., the typical causes of its tokenings). However, for the word to have any determinate meaning (i.e., to have a condition for correct application), the possibility of genuine error/misapplication should be made available to the speaker. Davidson thinks whether the speaker's particular application of a word is relevantly similar to her past applications (i.e., whether the application is correct) should not be something that the speaker herself could answer from the solipsist perspective. There should be another person observing the speaker's reactions to objects in the environment, who can judge their correctness from a different perspective than the speaker's own. Because meaning, in basic cases, can only emerge through this kind of interpersonal triangulation, Davidson's view of meaning-determination (or, constitution) is in an important way social, although not communitarian. From this view, it follows that there could not be such a thing as a non-shareable solitary language and whatever a speaker can come to mean by using a certain symbol should be *interpretable* by another person. As I will explain in Section 5, recognizing this consequence of Davidson's fuller account of meaning-determination is key to dissolving the worry I raised for Davidson in Section 3.

5. Davidson on the Social Dimension of First-Person Authority

The discussion in this section proceeds in two steps. In the first subsection, I will argue that the discussion in Section 4 allows Davidson to respond to the worry formulated in Section 3. To give more support to this claim, in the second subsection, I will explain how it also allows us to interpret some of the key passages from "First-person authority" and "Knowing One's Own Mind," where Davidson clearly hints at, but does not really elaborate on, the social aspect of his account of meaning-determination.

5.1. Davidson's response to the skeptical worry

The worry that I raised for Davidson's official account in Section 3 was this: on Davidson's account, a speaker's immunity to regular linguistic errors is explained by his non-communitarian semantic externalism,

according to which the meanings of words in a particular speaker's language are determined generally by the speaker's usage. I explained how this might seem to be in tension with the ordinary view that, for a speaker, the point of making a mental self-ascription should be to allow a hearer to acquire knowledge about the speaker's current mental state by taking her self-ascriptive utterance at face value. Davidson's acceptance of "a fairly extreme form of individualism about meaning," (Davidson 1993a, 250) might suggest that he admits the possibility of non-sharable solitary language. This can make room for the possibility that a hearer, no matter what evidence she may manage to collect about the speaker's language use, will never be able to articulate the content of a speaker's belief correctly, using her own words. This in turn seems to make Davidson's account of first-person authority subject to the skeptical worry concerning knowledge about other minds: even if Davidson's account can explain why there is a presumption that a speaker's self-ascriptive utterance is correct, it does so in a way that makes it unintelligible how this can also allow a hearer to acquire knowledge about the speaker's current state of mind by taking the speaker's self-ascriptive utterance at face value. If so, his view would be accused of failing to incorporate the social dimension of first-person authority.

Now, given the discussion in the previous section, we can understand why Davidson's account is not susceptible to this kind of worry. Recall that the key component in Davidson's account of first-person authority is the idea that a speaker cannot regularly misuse her own words because it is the speaker's use that confers meanings on them in the first place. Also, Davidson rejects the Burgean communitarian externalist view in part because he thinks it compromises first-person authority. His account of meaning-determination is thus individualist: a speaker need not use words in the same way as others do. However, we can now see that this captures merely one side of his account of meaning-determination, because he also argues, thorough the triangulation argument, that there can be nothing that a speaker can mean that is not in principle interpretable by another. It is crucial for a word to have any determinate meaning that some person other than the speaker herself could grasp its condition for correct application.

This allows us to resolve the problem in Davidson's framework. Although each speaker is, as Davidson argues, authoritative in determining what her words mean, there is also a crucial social constraint: what they can mean should be in principle interpretable to others. Therefore, although there is a crucial asymmetry between a speaker and a hearer concerning their knowledge about the meanings of the speaker's words, the asymmetry should not be taken to imply that there could be some content that is absolutely private to the speaker in the sense that others could never be able to articulate what it is. The possibility of such absolute privacy is ruled out by Davidson's triangulation argument that meaning can emerge only if there are two people interpreting one another's words in the shared environment.

When Davidson's account of first-person authority is properly integrated with his triangulation argument, we can see that his overall account is actually tailored to providing a plausible response to the skepticism concerning knowledge about other minds and thereby capturing the social dimension of first-person authority. Davidson, at least in one place, seems to note this point himself. The following passage comes after his brief discussion of triangulation in "Epistemology Externalized":

> Knowledge of one's own mind is personal. But what individuates that state at the same time makes it accessible to others, for the state is individuated by causal interplay among three elements: the thinker, others with whom he communicates, and an objective world they know they share. (Davidson 1990, 203–4)

In this passage, he suggests that there should be no real tension between his account of first-person authority and the ordinary view that the content of any belief that a speaker attributes to the self should be accessible to others. He briefly notes that this connects with his discussion on triangulation, but does not go on to explain how. Above, I offered a more thorough exposition of this point on behalf of Davidson by making the relevant connections more explicit.

5.2. Some suggestive passages from "First-person Authority" and "Knowing One's Own Mind"

As I noted earlier, even in "First-person authority" and "Knowing One's Own Mind," we can find some passages where Davidson clearly hints

at the idea that meaning is essentially social in the way that I explained above. Here is one such passage from "Knowing One's Own Mind":

> Though I reject Burge's insistence that we are bound to give a person's words the meaning they have in his linguistic community, and to interpret his propositional attitudes on the same basis, I think there is a somewhat different, but very important, sense in which social factors do control what a speaker can mean by his words. If a speaker wishes to be understood, he must intend his words to be interpreted in a certain way, and so must intend to provide his audience with the clues they need to arrive at the intended interpretation. [...] It is the requirement of learnability, interpretability, that provides the irreducible social factor, and that shows why someone can't mean something by his words that can't be correctly deciphered by another. (Davidson 1987, 28)

This passage clearly presents the idea that meaning is essentially social, and it has something to do with interpretability. Towards the end of the paper, he also makes a similar note that "[f]irst person authority, the social character of language, and the external determinants of thought and meaning go naturally together" (Davidson 1987, 38). The problem, again, is that he does not go on to explain exactly *how* they go together.

The discussion in the previous section can also help us understand some of the passages that may seem to pose a problem for commentators such as Jacobsen or Child, who endorse the standard interpretation that Davidson's account of first-person authority rests on his theory of meaning-determination. For example, in the last paragraph of "First-person Authority," Davidson notes this:

> There is a presumption—an unavoidable presumption built into the nature of interpretation—that the speaker usually knows what he means. So there is a presumption that if he knows that he holds a sentence true, he knows what he believes. (Davidson 1984, 14)

Here, Davidson notes that his account of first-person authority results from the "nature of interpretation," which may seem to implicate that it *does not* come from his account of meaning-determination. In fact, there is at least one commentator who criticizes the Jacobsen-Child line of interpretation by accusing them of ignoring this passage (Hossein Khani 2021, 6.1). However, we can see why Davidson's remark here need not pose any trouble for the standard interpretation. In his system, the issues of interpretation and meaning-determination are in fact inseparable.

For this reason, if his account of first-person authority is based on his view of meaning-determination, it must be *also* based on his view on the nature of interpretation, at least indirectly.

In light of this, we can also note that it is perhaps misleading, although not mistaken, to claim that in Davidson's framework, "the fundamental justification for the presumption that speakers generally know what their words mean comes from considerations about meaning-determination" and "it does not come directly from considerations about the *process* or *procedure* of interpretation" (Child 2007, 165). This is misleading because it can mask the crucial point that the social constraint of interpretability is also a key component in Davidson's account of meaning-determination after all. As I explained, this point is not fully transparent in Davidson's discussion of first-person authority. My claim has been that failing to recognize this point can lead to the mistaken idea that Davidson's account of first-person authority fails to recognize the social dimension of first-person authority. Showing why this is not the case takes the kind of interpretative work that I have undertaken in this paper.

6. Conclusion

The aim of this paper has been to show how Davidson's account of first-person authority, his rejection of the Burgean communitarian view of meaning-determination, and his view on the essentially social nature of meaning are connected in several important places. Specifically, if one is misled to think that Davidson subscribes to the view of meaning that allows for the possibility of solitary language, it would seem that his account of first-person authority fails to recognize the social dimension of first-person authority and even becomes susceptible to the skeptical worry concerning knowledge about other minds. To understand Davidson's answer to the problem, we need to recognize how his rejection of the Burgean communitarian idea and his insistence on the essential sociality of meaning ultimately hang together. This also allows us to interpret some suggestive remarks from "First-person Authority" and "Knowing One's Own Mind" where Davidson hints at, but does not go on to fully explain, connections between the issues of first-person authority and interpretability. As a result, we are now in a better position to see how Davidson's theoretical system as a whole is indeed tailored

to capturing the social dimension of first-person authority. First-person authority concerns not just how a speaker can usually make true mental self-ascriptions, but also how a hearer can accept them at face value and thereby gain knowledge about the speaker's mind. Although only the first, personal, dimension has been the main focus of the discussion in the literature, in Davidson's original discussion, we can find interesting theoretical resources to address both personal and social dimensions of the phenomenon in a unified framework. My aim has been to establish this point carefully to show that Davidson's discussion is worth reexamination in light of the recent discussion that highlights the essentially social dimension of first-person authority.

Acknowledgements

References

Bar-On, Dorit. 2004. *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Oxford University Press.

Beisecker, David. 2003. "Interpretation and First-Person Authority: Davidson on Self-Knowledge." *Southwest Philosophy Review* 19(1): 89–96.

Bernecker, Sven. 2013. "Triangular Externalism." In *A Companion to Donald Davidson*, edited by E. Lepore and K. Ludwig, 443–45. Chichester, West Sussex: Wiley-Blackwell.

Borgoni, Cristina. 2019. "Authority and Attribution: The Case of Epistemic Injustice in Self-Knowledge." *Philosophia* 47(2): 293–301.

Burge, Tyler. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4(1): 73–122.

———. n.d. *Cognition Through Understanding: Self-Knowledge, Interlocution, Reasoning, Reflection*. Oxford: Oxford University Press.

Child, William. 2007. "Davidson on First Person Authority and Knowledge of Meaning." *Nous* 41(2): 157–77.

———. 2013. "First-Person Authority." In *A Companion to Donald Davidson*, edited by E. Lepore and K. Ludwig, 533–49. Chichester, West Sussex: Wiley-Blackwell.

———. 2017. "Wittgenstein and Davidson on First-Person Authority and the Univocality of Mental Terms." In *Wittgenstein and Davidson on Language, Thought, and Action*, edited by C. Verheggen, 159–85. Cambridge: Cambridge University Press.

Coliva, Annalisa. 2016. *The Varieties of Self-Knowledge*. London: Palgrave Macmillan UK.

———. 2019. "Replies to Commentators." *Philosophia* 47(2): 343–52.

Davidson, Donald. 1984. "First Person Authority." *Dialectica* 38(2-3): 101–11. Reprinted in Davidson 2001a.

———. 1986. "A Nice Derangement of Epitaphs." In *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, edited by E. Lepore, 433–46. New York: Blackwell.

———. 1987. "Knowing One's Own Mind." *Proceedings and Addresses of the American Philosophical Association* 60(3): 441–58. Reprinted in Davidson 2001a.

———. 1990. "Epistemology Externalized." *Dialectica* 45(2-3): 191–202. Reprinted in Davidson 2001a.

———. 1991. "Three Varieties of Knowledge." In *A. J. Ayer: Memorial Essays*, edited by A. Phillips Griffiths. New York: Cambridge University Press. Reprinted in Davidson 2001a.

———. 1992. "The Second Person." *Midwest Studies in Philosophy* 17(1): 255-267. Reprinted in Davidson 2001b: 107–22.

———. 1993a. "Reply to Bernhard Thöle." In *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers (Foundations of Communication)*, edited by R. Stoecker, 248–50. Hawthorne: De Gruyter.

———. 1993b. "Reply to Eva Picardi." In *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers (Foundations of Communication)*, edited by R. Stoecker, 210–12. Hawthorne: De Gruyter.

———. 1994. "The Social Aspect of Language." In *The Philosophy of Michael Dummett*, edited by B. McGuinness and G. Oliveri, 1–16. Reprinted in Davidson 2005. Dordrecht: Springer.

———. 1997. "The Emergence of Thought." *Erkenntnis* 51(1): 7–17. Reprinted in Davidson 2001b: 123–34.

———. 2001a. "Externalisms." In *Interpreting Davidson*, edited by P. Kotatko, P. Pagin, and G. Segal, 1–16. Stanford: CSLI Press.

———. 2001b. *Subjective, Intersubjective, Objective*. Oxford: Oxford University Press.

———. 2005. *Truth, Language, and History*. Oxford: Clarendon Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and Ethics of Knowing*. New York: Oxford University Press.

Goldberg, Sanford C. 2004. "Radical Interpretation, Understanding, and the Testimonial Transmission of Knowledge." *Synthese* 138(3): 387–416.

Hacker, P.M.S. 1997. "Davidson on First-Person Authority." *Philosophical Quarterly* 47(188): 285–304.

Higginbotham, James. 1998. "On Knowing One's Own Language." In *On Knowing Our Own Minds*, edited by C. Wright, B. C. Smith, and C. Macdonald, 429–41. Oxford: Clarendon Press.

Hossein Khani, Ali. 2021. "Davidson on Self-Knowledge: A Transcendental Explanation." *The Southern Journal of Philosophy* 59(2): 153–84.

Jacobsen, Rockney. 2009. "Davidson and First-Person Authority: Parataxis and Self-Expression." *Pacific Philosophical Quarterly* 90(2): 251–66.

Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language*. Cambridge, MA: Harvard University Press.

Ludwig, Kirk A. 1994. "First-Person Knowledge and Authority." In *Language Mind and Epistemology: On Donald Davidson's Philosophy*, edited by G. Preyer, 367–98. Dordrecht: Kluwer Academic Publishers.

Macdonald, Cynthia. 1995. "Externalism and First-Person Authority." *Synthese* 104: 99–122.

Myers, Robert H., and Claudine Verheggen. 2016. *Donald's Davidson Triangulation Argument: A Philosophical Inquiry*. New York: Routledge.

Picardi, Eva. 1993. "First-Person Authority and Radical Interpretation." In *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers (Foundations of Communication)*, edited by R. Stoecker, 197–209. Hawthorne: De Gruyter.

Schwengerer, Lukas. 2021. "Beliefs over Avowals: Setting Up the Discourse on Self-Knowledge." *Episteme* 18(1): 66–81.

Smith, Barry C. 1998. "On Knowing One's Own Language." In *On Knowing Our Own Minds*, edited by C. Wright, B. C. Smith, and C. Macdonald, 391–428. Oxford: Clarendon Press.

Stanley, Jason, and Timothy Willlamson. 2001. "Knowing How." *Journal of Philosophy* 98(8): 411–44.

Sultanescu, Olivia, and Claudine Verheggen. 2019. "Davidson's Answer to Kripke's Sceptic." *Journal for the History of Analytical Philosophy* 7(2): 9–28.

Tanaka, Ryo. 2024. "Semantic Externalism, Self-Knowledge, and Inter-Personal Mental Ascription: A Neglected Puzzle." *Revista Portuguesa de Filosofia* 80(3): 819–40.

Thöle, Bernard. 1993. "The Explanation of First Person Authority." In *Reflecting Davidson: Donald Davidson Responding to an International Forum of Philosophers (Foundations of Communication)*, edited by R. Stoecker, 213–47. Hawthorne: De Gruyter.

Verheggen, Claudine. 2017. "Davidson's Semantic Externalism." *Argumenta* 3(1): 145–61.

Winokur, Benjamin. 2021. "Davidson, First-Person Authority, and Direct Self-Knowledge." *Synthese* 199(5–6): 13421–40.

———. 2023. "Authority as (Qualified) Indubitability." *Inquiry* 25.

Wright, Crispin. 2001. *Rails to Infinity*. Cambridge, MA: Harvard University Press.

Yalowitz, Steven. 1999. "Davidson's Social Externalism." *Philosophia* 27(1-2): 99–136.

Journal for the History of Analytical Philosophy

jhaponline.org