

REFERENCE WITHOUT DEFERENCE

Herman Cappelen and Max Deutsch

Forthcoming in Una Stojnić and Ernie Lepore (Eds) *The Oxford Handbook of Contemporary Philosophy of Language*

On a standard understanding of externalist metasemantic theories, such theories require a speaker to defer to other speakers in order to share content with them. We argue that this standard understanding is mistaken, and that, on a proper understanding of externalism, sharing content does not depend in any way on deference, either to experts, or one's linguistic community. We defend a version of externalism that we call 'pure externalism', and we argue that the idea that shared content requires deference is a residue of internalism/descriptivism to which externalists ought to be opposed. We also argue that, despite common belief to the contrary, several of the originators of externalism, including Tyler Burge, do not think that deference is in any way relevant, let alone required, for shared externalist content.

reference; deference; meaning; externalism; linguistic community; experts

1. Introduction

One (or the) core insight from the externalist tradition in metasemantics is that speakers can talk about the same object (or the same kind of thing) even if they have very different views about that object (or kind). Shared reference and shared meanings are not established through shared beliefs or a common creed but, instead, through a form of causal or historical interconnectedness. Traditional externalism, however, incorporates a cluster of ideas that are in tension with the renunciation of a requirement of a common creed. One of these ideas, which the tradition has inherited from Burge, Putnam, and Evans, is the idea that *deference to experts* plays an important role in a semantic theory. Here's one way to articulate the received view (this is an initial approximation, we'll consider precisifications below):

Expert deference (ED): For some terms in natural language, non-expert speakers must defer to experts in order to be competent users. They share contents (with each other and the experts) through shared deference to experts.

We say that the tradition inherited versions of ED from Putnam, Burge and Evans, but it's a tainted inheritance because (as we point out below) those authors didn't argue that expert deference was a semantic phenomenon (they didn't even talk about expert deference). They have been misrepresented. We go on to argue that they were right to not endorse ED, because it's false. Externalism should be purged of deference: there can be shared *reference* without *deference*.

The Kripke of *Naming and Necessity* was not a proponent of ED, but he did advocate a related view, which we'll call Communal Deference:¹

Communal deference (CD): In order to use a term in a language to refer to what it refers to in that language, a speaker has to intend to preserve reference—they have to intend to use it with the meaning it has in that language.

What ED and CD have in common is that they require that a certain cognitive state—a state of deference—be shared by those who share meanings or use terms to speak of the same referents. According to ED, non-expert speakers of English, for example, need to all be in the state of deferring to experts on the extension of 'arthritis' in order to share in the communal use of that term. According to CD, speakers (who are not introducing a term) must all have an intention of a certain form: the intention to use the term with the same meaning and reference as other speakers in the linguistic community.

¹ This is an initial approximation; more precise versions will be discussed below.

Our aim in what follows is to show that the externalist tradition should never have endorsed (any version of) ED or CD. These are, in effect, internalist requirements that have no place in an externalist framework. We argue for three central claims:

- (i) ED and CD are both unmotivated.
- (ii) ED and CD are subject to counter-examples of the same kind as those that externalists use against internalism more generally. It is close to inconsistent for externalists to reject this data while endorsing the standard arguments for externalism.
- (iii) Externalism without deference—what we call ‘*pure externalism*’—can stand on its own: externalism isn’t strengthened by appeal to deference.

We start with a discussion of ED; then we turn to CD.

2. No Deference to Experts in Putnam or Burge

Our goal is to show that expert deference should play no role in externalist metasemantics, but before we turn to that, we want to do some history of philosophy. Two of the philosophers most closely associated with the idea of expert deference, Hilary Putnam and Tyler Burge, didn’t endorse any version of ED. They didn’t even talk about deference. Putnam, who introduced the notion of the Division of Linguistic Labor, says, in a typical passage,

[E]veryone to whom gold is important for any reason has to acquire the word “gold”; but he does not have to acquire the method of recognizing if something is or is not gold. *He can rely on a special subclass of speakers.* The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing if something is in the extension (“criteria”), etc.—*are all present in the linguistic community considered as a collective body; but that*

collective body divides the “labor” of knowing and employing these various parts of the “meaning” of “gold.” This division of linguistic labor rests upon and presupposes the division of non-linguistic labor, of course.

(Putnam 1975: 145)

In this and related passages, Putnam makes a couple of points:

- (1) To be a competent user of ‘gold’, you don’t need to be able to recognize gold.
- (2) The ability to recognize gold—distinguish gold from not-gold—is present in the linguistic community as a whole.

Neither (1) nor (2) provide support for the view that deference to experts is a metasemantic phenomenon. We take (1) to be as close to a non-controversial claim as we can have in this domain. It is certainly true that recognitional abilities are distributed unevenly across the linguistic community and that you can be a competent user of a term even if you don’t know how to distinguish what’s in its extension from what’s not. That has nothing to do with deference to experts.² (2) is a claim about what recognitional capacities we should expect to find in a linguistic community. We think it is an implausible claim (we could use ‘COVID’ before people knew how to recognize all instances of COVID, for example), but that’s not our focus here. The important point is that it’s not a claim about deference. Even if it is true that some speakers must have special recognitional capacities with respect to the

² We suspect that it is Putnam’s claim that we who can’t distinguish all instances of gold from non-gold can nevertheless “rely on a special subclass of speakers” that led to the misinterpretation (of Putnam as a defender of ED). Interpreted correctly, however, Putnam is saying merely that in order to *tell* or *know* that something is gold, we might have to rely on others. This *epistemic* form of reliance is entirely routine and has no connection to metasemantics. We say more about *epistemic* versus *semantic* deference later.

denotation of ‘gold’, nothing follows about how non-expert speakers must stand in the deference relation to those people. The word ‘deference’ doesn’t occur in “The Meaning of ‘Meaning’” and the paper has nothing to say on the topic.

‘Deference’ and cognate terms are absent also from Burge’s ‘Individualism and the Mental’ (Burge 1979). Burge, in that paper, does not claim that speakers must defer to experts on ‘arthritis’ in order to be competent users of ‘arthritis’. The relevant part of Burge’s ‘arthritis’ thought-experiment is the second stage.³ Here, Burge asks us to consider a counterfactual situation in which exactly the same things (non-intentionally described) happen to the patient who thinks he has arthritis in his thigh, but where “physicians, lexicographers, and informed laymen apply ‘arthritis’ not only to arthritis but to various other rheumatoid ailments. The standard use of the term is to be conceived to encompass the patient’s actual misuse.” In sum,

The person might have had the same physical history and non-intentional mental phenomena while the word ‘arthritis’ was conventionally applied, and defined to apply, to various rheumatoid ailments, including the one in the person’s thigh, as well as to arthritis.

(Burge 1979: 78)

Note that Burge simply *stipulates* that, in the counterfactual situation, the term ‘arthritis’ doesn’t denote arthritis but some other rheumatoid condition. He doesn’t tell us much about how that has happened, but he mentions conventions and definitions. Burge says that in the counterfactual situation, “the patient lacks some—probably all—of the attitudes commonly attributed with content clauses containing ‘arthritis’ in oblique occurrence.” As a result, “the patient’s mental contents differ while his entire physical and non-intentional mental histories, considered in isolation from their social context, remain the same.”

Bracket for a moment the question of whether the argument is sound. Our interest is in whether the notion of expert deference plays a role in it. Nothing in this argument has

³ For those not familiar with this thought-experiment, see Burge (1979).

specifically to do with the presence of an expert community. The central move is to stipulate that a lexical item—‘arthritis’—has a different meaning in the linguistic community and that this doesn’t need to affect an individual speaker. An individual speaker’s physical and non-intentional mental history could remain the same even if an expression she’s competent with has a different meaning. Proper names would provide an equally good example: consider Alfred, to whom Max can refer using ‘Alfred’ even though he knows very little about him other than that he’s a doorman at a bar in Hoboken. Next, imagine a counterfactual situation in which Alfred’s identical twin, Albert, was named ‘Alfred’ and ended up doing all the things that Alfred did, including the stint as doorman in Hoboken. The name ‘Alfred’ could have reached Max just as it has in the actual world, while Max’s entire physical and non-intentional mental history would remain the same. He would have different mental states because he would be thinking and talking about Albert (using ‘Alfred’), not Alfred. The difference between this argument and the one Burge runs using ‘arthritis’ has to do with semantic and syntactic category, not with deference to experts. In the Burge example, the people introducing the term are doctors and people in the medical community, but that’s not an essential part of the example. The power to introduce new terms with new meanings isn’t distinctive of a group called ‘experts’. Alfred’s parents had the power to call him ‘Alfred’. A group of people was able to introduce the term ‘woke’—we don’t know who they are, and the mechanism is poorly understood, but somehow it happened. You can ask general questions about the who–how–when of linguistic creativity, but the answers to those questions are not what Burge’s argument depends on. In short: the essential elements of Burge’s argument is independent of both deference and experts.

Despite Putnam and Burge making no appeals to expert deference in their famous externalist papers, the view that they argued that expert deference is a requirement on shared

externalist meaning and reference is very widespread. Here is a representative example from Curtis Brown (2022) in the *Stanford Encyclopedia of Philosophy*:

Thus, what we are thinking about depends not only on our intrinsic properties, but also on expert opinion. We *defer* to the experts with regard to what exactly we are thinking about. For this reason, this sort of contribution of the social environment is sometimes referred to as “semantic deference.” The phenomenon of semantic deference is closely related to what Putnam memorably termed “the linguistic division of labor.” Putnam’s idea was that as long as there are experts on what certain words refer to, we do not all need to have that specialized knowledge; we can rely on the knowledge of the experts. In Burge’s treatment, however, the phenomenon is not merely linguistic: it is not just that we defer to the experts on the meaning of the *word* ‘arthritis’; we also defer to the experts on the nature of the *disease* arthritis.

(Brown 2022: §2.2)

Our view is that this is simply a mistaken interpretation of Putnam and Burge. As we pointed out above, Putnam and Burge don’t mention deference, and their externalism does not require anything describable as ‘deference to experts’ for shared, communal meaning or reference.

3. Against Semantic Deference to Experts

For historical accuracy, it matters what Putnam and Burge said, but maybe they just failed to realize the semantic importance of expert deference. Maybe the interpreters of these texts added a valuable element. Unfortunately, we don’t share this optimistic view. The problem is not just that arguments for the semantic significance of expert deference are lacking in the papers that spawned the externalist tradition. The more important problem is that there are good arguments against the view and no evidence in its favor.

3.1 Identifying ‘deference’: Epistemic deference vs semantic deference

Before we make our critical points here, first some common ground. Epistemic deference to experts should be entirely non-controversial. For various theoretical purposes, there will be many ways to precisify ‘experts’ and ‘deference’, but working with a pre-theoretical notion, it’s hard to deny that if you want to learn about arthritis, you should defer to people who know a lot about arthritis and who are honest and constructive interlocutors. Much of what we learn about the world we learn through others, and it requires some kind of deference (in a loose, non-theoretical, sense of that term). That, however, is a view you can hold independently of your view of how meaning is grounded and determined. This kind of non-controversial epistemic deference, on its own, doesn’t amount to any kind of semantic or metasemantic thesis.

Those who advocate for the semantic significance of deference to experts need to distinguish that phenomenon from the mundane epistemic kind of deference to people who are knowledgeable about some subject. One strategy for that is to distinguish two kinds of deference:

- *epistemic deference* has, roughly, the form: ‘I will assign high credence to what the experts tell me about T’;
- *semantic deference* has, roughly, the form: ‘I want to refer to whatever the experts refer to when they use “T”.’

Proponents of the semantic significance of deference need not appeal to semantic deference. They could, for example, appeal to a distinctive pattern of epistemic deference. We’ll explore both of these options below, but we’ll start with pointing out the often overlooked fact that expertise on T typically doesn’t go hand-in-hand with competency with ‘T’.

3.2 The discrepancy between T-expertise and ‘T’-expertise

Identifying expertise and experts is difficult for many reasons, but we want to focus on a distinctive and often overlooked problem: the experts on a topic—say, experts on arthritis—need not be English speakers and so need not have any competence with the term ‘arthritis’.

Consider three options for identifying the relevant experts:

Object-level experts: this is the group of people who are experts on arthritis. Note that there’s no requirement on membership in this group that members have used or heard the word ‘arthritis’. Experts on arthritis who are monolingual speakers of Mandarin, for example, have no ability to use ‘arthritis’ (they do, of course, have the ability to use ‘关节炎’ and to think about what it means, namely the condition afflicting people’s joints we call ‘arthritis’). More specifically, their expertise has no bearing on the use of that English term.

Meta-linguistic expert: we can imagine an expert on words, maybe some kind of linguist, who knows a lot about, for example, the word ‘arthritis’ (maybe about its etymology, phonology, syntax, etc.). Such an expert, however, need not have any medical expertise and might be as confused about what arthritis is as any layperson.

Hybrid expert: here is our best effort of a rational reconstruction of a hybrid of the two previous versions: an expert on arthritis (say, a GP who learned about arthritis in medical school) and a regular English speaker who has ‘arthritis’ in their vocabulary.

In sum, a pretty normal doctor.

If what the word ‘arthritis’ means when used by non-experts is, in part, determined by the speaker’s deference to experts, the relevant notion of ‘expert’ should be one of those three. But all three are problematic:

- The meta-linguistic expert is out: proponents of ED are not suggesting that we all should be constantly deferring to linguists.

- It would also be surprising if it was the object-level expert we had to defer to. How can people who have never spoken English influence the meaning of an English word? These speakers are deserving of epistemic deference (more on that below), but it's hard to make sense of the idea that we English speakers defer semantically to them.
- So proponents of ED will need the hybrid expert. However, as presented, 'hybrid expert' can't be right. First, 'hybrid expert' is Anglocentric in an absurd way. It can't be true that Chinese speakers should semantically defer to object-level experts who speak English. The more general version would have to relativize semantic deference to the deferrer: "for speakers of language N , experts are those who speak N and have object-level expertise." However, and this is the second problem: there seems to be too weak a connection between object-level expertise and the relevant linguistic practice. To see what we have in mind, consider someone who is an expert on arthritis and speaks English as a second language and Mandarin as a first language. She might be a world leading expert on arthritis, but at the same time have a little gap in her English competency that leads her to misunderstand the meaning of 'arthritis' (and thinks it applies to rheumatoid conditions beyond the joints). Again, it's hard to see why this person should have a distinctive semantic status for English speakers. This point can be pushed further: it's not an important feature of the example that the imagined object-level expert has English only as a second language. Object-level competency is separable from linguistic competence: the world leading English-speaking experts on arthritis might all be using a Latin term for arthritis and be sloppy users of 'arthritis'. What proponents of ED really have in mind, we suspect, is experts on arthritis who are using the relevant terminology *correctly*—with the meaning it has in English.

Then we can't, at the same time, have the experts be the ones determining correct usage of the English expression.

We don't know how to solve the problem just sketched; that is, we don't know how those who think that deference to experts is an important semantic phenomenon should identify the relevant set of experts. In the objections below, we'll move between the three options, appealing to the one most favorable to proponents of ED.

3.3 Kripke-style arguments against ED

We'll first address the necessity version of ED (speakers share content only if they defer) and then briefly consider the even less plausible sufficiency claim (speakers share content if they defer). For those who accept Kripke's original "arguments from ignorance and error," it should be a short step to a rejection of deference to experts as a requirement on the sharing of meaning and reference. A central theme of *Naming and Necessity* (Kripke 1980) is that those who share reference don't require a common creed: there are, for example, no beliefs we need to share in order to all use 'Gödel' to denote Gödel. Kripke made us aware of that by pointing to cases where reference to an object was achieved despite lack of relevant beliefs (either because there were too few beliefs or the relevant beliefs were false). The same kind of argument brings out the semantic non-necessity of deference to experts. Consider the following scenarios:

1. *Wrong about the identity of the experts:* there are competent non-expert users of 'T' who want to defer to T-experts but are wrong about who the experts are. These speakers have the correct view of what deference is, and what an expert is, but fail to pick out the right experts. Let's say the correct theory of expertise is ET (since it's very difficult to spell out ET, we'll remain content with a placeholder). We can imagine that many of those who reject the claims of 'western medicine' have the right view of expertise. Insofar as any of us are

operating with ET, they are too. Their problem is that they fail to pick out those who satisfy ET. They might, for example, think that certain religious gurus, not doctors, are experts on arthritis (maybe because they think the gurus have more knowledge and better methodology (mystical insight)). As a result, they fail to defer to experts (since that means deferring to those who satisfy ET).⁴ This failure, however, doesn't imply that they fail to use 'arthritis' to denote arthritis.⁵

2. *Wrong about what deference is*: the speakers we imagined above have our notion of deference but fail to identify the experts. Now, consider those who correctly identify the experts but are wrong about what deference is. Some cases to consider:

- (i) S thinks that in order to defer to someone's view about a subject matter, they have to endorse that person's view (about the subject matter in question).

⁴ A difficult-to-distinguish version of this involves speakers who have a false view about what experts are. They don't accept ET (in whatever way 'normal' speakers are alleged to do). They might explicitly endorse a false view of expertise, FET. They are, however, good at picking out those who satisfy FET, and they defer to them. As a result of their mistaken assumption about what constitutes expertise, they end up being mistaken about who the experts are.

⁵ There's an analogous point to be made about a community-wide failure to identify the right experts. That's probably the situation we are in: more people in the world defer to non-experts than defer to experts. If we add together all the religious gurus and bullshitters in various domains, it's not implausible that they get broader popular support than real experts.

- (ii) S1 thinks that in order to defer to someone's view about a subject matter, they have to assign minimally a credence to that person's view (about the subject matter in question).
- (iii) S2 thinks that in order to defer to someone's view about a subject matter, it suffices to say 'yes' to whatever that person says (about the subject matter in question).

These views about deference can't all be right. As a matter of fact, we (the authors of this chapter), don't know what deference really is. For the sake of argument, assume deference (as understood by proponents of ED) is D. Our point, here, is that competent users of 'arthritis' can be wrong about D and be disposed to act in non-D-like ways.

3. *No experts*: in some cases, *there simply are no relevant experts*, so no experts to whom to defer. If deference to experts were a necessary condition, the result would be a linguistic and communicative breakdown. That's not the case.

First, there are domains where, whatever expertise might be, it's very implausible to assume that there could be any. Take, for example, tallness (the property of being tall relative to something or other). There are no experts on that, but people can talk about tallness, no problem. Second, consider domains where there are a few experts, and they might all disappear. One of the authors of this chapter (HC) was once on a flight with all the world's experts on quotation.⁶ If that plane had gone down, there would be no more experts on quotation, but people would still be able to use 'quotation' to denote quotation.

⁶ Assuming, for the sake of argument, that there can be experts on quotation (if there are no such experts, then the first problem discussed—the absence of experts—will apply broadly).

4. *Refusal to defer to experts*: finally, consider speakers who have a correct view both of what experts are and of what deference is. However, they might, for example, think that experts are cogs in a social machinery of repression and injustice, and as a result, they refuse to defer to them. They think the practice of expert deference is socially destructive and that the way to achieve knowledge is some generalized version of Wikipedia where everyone is equal and ‘expertise’ is not recognized. The details of this can vary a lot, and the point is that no matter how the details are worked out, this kind of rebellious attitude doesn’t prevent someone from using ‘arthritis’ to denote arthritis.

If you’re someone who agrees with the general line of anti-descriptivist, pro-externalist argumentation in *Naming and Necessity*, these cases will seem natural and continuous with that argumentation. People with all kinds of weird propensities and false beliefs can share content with us. That extends to their beliefs about deference, and dispositions to defer.

3.4 Incomplete understanding and deference in Burge

Some authors distinguish between complete and incomplete understanding of a term (Recanati 1997, 2000; Wikforss 2017). These authors might join our rejection of deference to experts but, instead, insist that those who have incomplete understanding defer to those whose understanding is complete. The view would go along these lines:

Necessity of deference to completers

- (i) Among competent users of a term, T, some of them have an incomplete understanding (call these ‘the T-incompleters’) and others have a (more) complete understanding (call these ‘the T-completers’)
- (ii) The T-incompleters must defer to the T-completers (in order to be competent users of T).

The objections to the necessity version of ED also undermine this argument. Just replace ‘expert’ with ‘completer’ in section 2. Competent incompleters can (a) be wrong about the identity of the completers; (b) they can be wrong about the nature of deference; (c) there might be no T-completers; and (d) a T-incompleter can refuse to defer to the T-completers.

It is also worth pointing out that the necessity-of-completers argument isn’t Burge’s and that externalists should reject the terminology of ‘complete vs incomplete understanding’. Burge doesn’t talk about people who have ‘complete understanding’ (the term or cognates don’t occur in his seminal paper), and the notion of incompleteness isn’t supposed to do much theoretical work for him. He says, “The notions of misconception, incomplete understanding, conceptual or linguistic error, and ordinary empirical error are to be taken as carrying little theoretical weight” (Burge 1979: 88). In Burge’s picture, there’s no sharp distinction between empirical and conceptual knowledge. There are some people who know a little more than others, but he doesn’t contrast the incompleters with the completers, and the idea of ‘completeness’ doesn’t make sense on his view.⁷

There are, of course, theorists other than Burge who think that there’s a set of competencies (dispositions or beliefs) that are such that possessing them makes your understanding ‘complete’. That would in effect require picking a set of beliefs as those that are meaning- (or competency-) constitutive. Again, our goal here isn’t to rehash the

⁷ He says, “I do not believe that understanding, in our examples, can be explicated as independent of empirical knowledge, or that the conceptual errors of our subjects are best seen as ‘purely’ mistakes about concepts and as involving no ‘admixture’ of error about ‘the world.’ With Quine, I find such talk about purity and mixture devoid of illumination or explanatory power” (Burge 1979: 88).

arguments against that kind of view but to point out that Kripkeans have already rejected this view. There's no such privileged cluster of beliefs (or dispositions) that makes your understanding 'complete'. That would be a coarse form of the kind of internalism that *Naming and Necessity* refutes. The arguments against a descriptive view of 'Godel' show both that those with 'incomplete' understanding of 'Godel' can nevertheless refer to Godel and that a complete understanding isn't what it takes to refer to Godel.

3.5 Against the sufficiency of deference to experts

Above, we discussed deference to experts as a necessary condition for shared reference and meaning between experts and non. We think the claim that deference to experts is *sufficient* for shared content is even less plausible and is also not typically what proponents of a deference condition have in mind (they have in mind the already-discussed necessary condition version). But we will, here, briefly explain why the sufficiency version is so implausible.

We assume that there is agreement on the general principle that just because a speaker wants (or intends) an expression, E, to mean M, that's insufficient for her utterances of E to express M. If Max wants 'dog' to mean the same as what 'cat' means, and utters 'I own a dog', his utterance doesn't express the content that Max owns a cat. That's now how language works. But now we have a reason to deny that deference to what someone else means by E is sufficient for it to mean what that someone else means by E. Suppose Bob means cat by 'dog'. Max can't get 'dog' to mean cat just by intending his utterance of 'dog' to mean the same as what Bob means by it. Meaning change isn't that easy. And it remains not that easy, even if Bob happens to know more about cats and dogs than Max does.

4. Deference to the Linguistic Community (or Standard Usage)

We turn now to a different form of deference. In *Naming and Necessity*, Kripke doesn't talk about deference to experts—that came, taintedly, from Putnam and Burge. Kripke does, however, talk about deference to *other speakers*:

When the name is 'passed from link to link', the receiver of the name must, I think, intend when he learns to use it with the same reference as the man from whom he heard it. If I hear the name 'Napoleon' and decide it would be a nice name for my pet aardvark, I do not satisfy this condition.

(1980: 96)

Many other authors have picked up on this remark, and it is now an established part of the received view. The requirement that speakers must intend to preserve reference in this way is emphasized by almost all authors in this tradition. To give just one example, here is a passage from Tye and Sainsbury:

Most public concepts have their referent before we acquire them. In acquiring them, we aim to use them with just the reference they already have in our community; this is part of what we understand by deference. Deference can be modeled (rather over-intellectually) as the recognition that others already use a concept, together with the desire to use the very same concept they use, with the very reference it has in their uses. It does not require the notion of an expert, nor does one need to think of "correct use" as use to form a true thought (though that is often the case), as opposed to use that is genuinely a use of the right concept. This picture makes very appealing the idea that conceptual reference is preserved by some kind of causal process relating uses of concepts, since change of reference would correspond to failure of an intention.

Note that Kripke and Sainsbury and Tye all connect deference to previous usage as a necessary condition, and they also claim that this is required in order to explain reference shifts. The reason Kripke brings up the pet aardvark is to emphasize that an established proper name like 'Napoleon' can change its reference from a French statesman to an African animal. He suggests that this change in reference can be explained by the renaming intention and that, if so, this provides evidence of the requirement that continuity of reference requires an intention to preserve reference.

In what follows, we aim to show, first, that an intention to preserve reference is not required for reference preservation. Second, we show that the rejection of this necessary condition is unrelated to the issue of reference shift: one could explain the shift in reference the way Kripke and Sainsbury and Tye do, even if one rejects, as we do, the view that an intention to preserve reference is necessary for reference preservation.

4.1 Against deference to standard usage

Our opponent here comes in many varieties, but for simplicity we will focus on the following versions:

Kripke's deference requirement (CD-K): for a speaker to be part of a communicative chain involving expression N, she has to intend, at the time when she learns it, to use N with the same referent as the person she learned it from.

Generalized deference requirement (CD-G): for a speaker of language L to use an expression N with the referent (or semantic value) it has in L, she has to intend (at some point) to use it with the referent it has in L.

As in our discussion of deference to experts, we think there are fairly straightforward counter-examples to these requirements and, moreover, they pattern with the paradigmatic arguments that Kripke uses in support of his view in *Naming and Necessity*:

N. Chomsky (NC), no reference: imagine a fictional character, N. Chomsky, modeled on the real person Noam Chomsky (but the real Chomsky is much more subtle than our imagined character). NC doesn't believe in reference. He thinks the whole idea is absurd and incoherent. NC has good arguments for this view, and he writes books about it. So when he learns a new word, he never has an intention of the form: I intend to use it to refer to the same object as the person I got it from. He never has an intention of the form I intend to use it with the same reference as other speakers in my linguistic community.

N. Chomsky, no language: NC denies not only the existence of reference but also the existence of languages, and he denies that linguistic communities play a role in an account of speech and interpretation. As a corollary, NC never has intentions expressible as 'I defer to those who speak the same language as me', nor expressible as 'I intend to use the term with the same meaning as other speakers of my linguistic community'.

Alfred the Anarchist: imagine a story beginning as some of the stories Kripke tells about reference fixing begin: the name 'Alpha' is given to a baby in a baptismal event. The name is then, as Kripke says, passed along as if in a chain. At some point, the chain reaches Alfred. Alfred has asked his neighbor why there's so much noise during the night. The neighbor answers, "We have a newborn kid, Alpha, and she often wakes up at night." Now, Alfred is an anarchist and he defers to nothing, he says. He has read Kripke's book *Naming and Necessity* but rejects Kripke's view: he will not intend to use the name with the same referent as the person he learned it from—"I'm a free spirit—and I mean what I want to mean," says Alfred. More specifically, he has a perfect understanding of Kripke's notion of deference, and he has trained himself not to have that intention when he hears a new name. Nonetheless,

he goes on to use the name ‘Alpha’, saying things like, “This kid, Alpha, makes a ruckus at night” and “I wish Alpha wouldn’t scream so much.”

These, we suggest, are cases where reference takes place despite CD-K and CD-G not being satisfied. These cases pattern with other Kripke-style arguments: there’s a creed (a belief, if you will) that a speaker fails to share with other speakers. The speakers have false beliefs (about who experts are or what deference or what the correct metasemantics is), but such false beliefs don’t undermine their ability to use English words with their standard meaning.

Let’s briefly consider one objection to the line of argument above: maybe what should be appealed to by proponents of CD-K and CD-G is an account of deference according to which beliefs and commitments about deference don’t determine whether an agent defers:

As long as Alfred behaves deferentially, we should ignore his beliefs, sayings, and commitments—real deference is revealed through actions.

We are then owed an account of what behavioral dispositions are constitutive of deference. We know of no such account, but suppose a proponent of b comes up with some behavioral dispositions, B, that constitute deference. We will tell Alfred about this (since Alfred is fictional, we can only do this fictionally). In response, Alfred decides to avoid B behavior. Whatever you suggest, we stipulate that Alfred, being the kind of anarchist he is, refuses to perform those actions. He will not B, under any circumstances, but he will perform alternative actions to compensate. Just how this will play out depends a lot on the details of B, but our conjecture is that, just as there is not cluster of beliefs that are necessary or sufficient for reference, there is no cluster of behavior that plays that role (and the dialectic will play out in parallel ways).

4.2 Kripke’s motivation: Reference shift

We turn now to Kripke’s motivation for introducing deference to the community. He introduces that idea when he talks about the possibility of reference shifts (and an objection

to his view that he attributes to Evans—that motivation is also mentioned in the passage from Sainsbury and Tye 2012). Talking about various examples of reference shift, Kripke says,

In all these cases, a present intention to refer to a given entity (or to refer fictionally) overrides the original intention to preserve reference in the historical chain of transmission. The matter deserves extended discussion. But the phenomenon is perhaps roughly explicable in terms of the predominantly social character of the use of proper names emphasized in the text: we use names to communicate with other speakers in a common language. This character dictates ordinarily that a speaker intend to use a name the same way as it was transmitted to him; but in the ‘Madagascar’ case this social character dictates that the present intention to refer to an island overrides the distant link to native usage. . . . To state all this with any precision undoubtedly requires more apparatus than I have developed here; in particular, we must distinguish a present intention to use a name for an object from a mere present belief that the object is the only one having a certain property, and clarify this distinction. I leave the problem for further work.

(Kripke 1980: 163)

4.3 Why reference shift doesn’t motivate CD

Two points about the appeal of CD as a step toward understanding reference shifts: (a) it’s not a very plausible view; (b) even if it were plausible, the proposed view wouldn’t motivate CD:

- (a) Let’s first put aside the question of whether Kripke’s suggestion can explain reference shift in some cases (for reasons given below, we doubt that). Even if Kripke were right and a renaming intention was needed to explain reference

shifts in *some* cases, that's compatible with the denial of the deference as a requirement on reference. Maybe—we're not committed to this—an overriding renaming intention, in some cases, can undermine semantic reference, but even so, there is no requirement that we have the positive intention to denote what the community denotes in order to denote what the community denotes. That's a structural point that combines nicely with the objections to CD that we've presented above.

- (b) Separately, the proposed explanation of reference shifts is implausible as a response to the kinds of cases that Evans discussed. On the face of it, the example Kripke and Evans appeal to—the change in reference of 'Madagascar'—doesn't involve the absence of deference intentions. This is not a case where the speakers at some point decided that they didn't want to refer to what the native speakers referred to. Marco Polo and pals were just confused about what native speakers referred to. Of course, Kripke's promissory note—"I leave the problem for further work"—could be doing all the work here. Again, Kripke's own ideas show how hard this additional work will be. Take a speaker who thinks she sees Smith when she really sees Jones, and says, 'Smith is smoking a pipe.' She intends to refer to 'Smith' (because she's using the name 'Smith' and knows that it denotes Smith), but she also wants to talk about the person she sees sitting over there, that is Jones. Kripke's (1977) view is that the semantic referent of 'Smith' hasn't changed; it still refers to Smith. The speaker referent—the object the speaker has in mind and wants to talk about—is Jones. So in this case, what the speaker wants to refer to doesn't affect the meaning of 'Smith' in the language. The relevant uses of 'Madagascar' would have to be different from that. Maybe the

idea is that, over time, with a sufficient number of speakers joining in, we get a gradual change in semantic referent. If some version of that is the view, then the point in (a) kicks in again: articulate this as the view that when there's a relevant kind of communal intention over time to denote D, then, under the right conditions, the semantic reference shifts to D. No need to also require an ongoing intention to defer.

4.4 Externalism without deference: Williamson on unity in diversity

We can think of the practice of using an expression with a common meaning as a shared coordinated activity among a group of people that can be spread over time and space. These speakers vary massively: they have different beliefs, dispositions, preferences, and stand in different relations to the non-linguistic world. Nonetheless, something enables them to co-denote and to share meanings over time. In *The Philosophy of Philosophy*, Timothy Williamson provides what we find an illuminating description of what creates unity out of diversity. Williamson says,

What binds together different events into the history of a single complex object, whether it be a stone, a tree, a table, a person, a society, a tradition, or a word? In brief, what makes unity out of diversity? Rarely is the answer to such questions the mutual similarity of the constituents. Almost never is it some invariant feature, shared by all the constituents and somehow prior to the complex whole itself—an indivisible soul or bare particular. Rather it is the complex interrelations of the constituents, above all, their causal interrelations. Although we should not expect a precise non-circular statement of necessary and sufficient conditions for the unity in terms of those complex interrelations, we have at least a rough idea of what it takes. The similarity of the constituents is neither necessary nor sufficient; different constituents can play

different but complementary roles in constituting the unity: both events in the head and events in the heart help constitute the life of a person.

(Williamson 2007: 125)

There are three core take-home message from this passage when applied to the question of how meaning is shared in a linguistic community:

1. On the negative side: content sharing doesn't depend on the speakers sharing properties. In particular, they need share no beliefs or attitudes. The anti-deference theme of this chapter is an instantiation of that general point.
2. *General anti-reductionist view*: it's a mistake to expect reductive (non-circular) necessary and sufficient conditions for what generates unity. This echoes some important remarks by Kripke in *Naming and Necessity*, where he says,

I'm always sympathetic to Bishop Butler's "Everything is what it is and not another thing"—in the nontrivial sense that philosophical analyses of some concept like reference, in completely different terms which make no mention of reference, are very apt to fail. Of course in any particular case when one is given an analysis one has to look at it and see whether it is true or false. One can't just cite this maxim to oneself and then turn the page. But more cautiously, I want to present a better picture without giving a set of necessary and sufficient conditions for reference.

(Kripke 1980: 94)

3. *The importance of causal interrelations*: Williamson, just as Kripke does, emphasizes the importance of casual interconnections. In this case, those are

interconnections between past speakers and current speakers and between current speakers.

Those who are reductionists will ask for a reductive account of these casual interconnections—that is in terms that eliminate ‘refer’, ‘mean’, ‘denote’, and cognates. That, however, is a theoretically implausible goal, not just for those trying to understand shared meaning but, more generally, for explaining unity in diversity. Note that, in saying this, we are not denying the importance of speakers’ attitudes, intentions, and beliefs. Those elements will play a role, but those roles cannot be laid out as neat necessary and sufficient conditions.

We want to end by pointing out that if you are already convinced by the arguments and views in *Naming and Necessity*, then abandoning deference, should in no way decrease your credence in the view. If you like the arguments against descriptive theories, they’re untouched by any of this. The positive picture—the communicative chain picture—is also unaffected. Recall that deference was introduced as a reply to an objection from Evans (involving reference change), and, as we showed, the reply to that objection still stands. Kripke just overreacted and went slightly overboard (or, at least, the interpretation of Kripke ended up being overboard). Now, for the non-Kripkeans, we don’t have that much to say: this chapter isn’t for them. Suffice it to say that whatever the source of their anti-Kripkean tendencies, the removal of deference shouldn’t add fuel to their fire.

References

- Brown, Curtis (2022). "Narrow mental content." In *Stanford Encyclopedia of Philosophy*.
- Burge, Tyler (1979). "Individualism and the mental." *Midwest Studies in Philosophy* 4(1): 73–122.
- Kripke, Saul A. (1977). "Speaker's reference and semantic reference." In Peter A. French, Theodore E. Uehling Jr, Howard K. Wettstein (eds), *Studies in the Philosophy of Language*. Minneapolis, MN: University of Minnesota Press, pp. 255–296.
- Kripke, Saul A. (1980). *Naming and Necessity: Lectures Given to the Princeton University Philosophy Colloquium*. Cambridge, MA: Harvard University Press.
- Putnam, Hillary (1975). "The meaning of 'meaning'." *Minnesota Studies in the Philosophy of Science* 7: 131–193.
- Recanati, Francois (1997). "Direct reference: From language to thought." *Noûs* 31(4): 538–556.
- Recanati, Francois (2000). "Deferential concepts: A response to Woodfield." *Mind and Language* 15(4): 452–464.
- Sainsbury, Richard Mark and Tye, Michael (2012). *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*. Oxford and New York: Oxford University Press.
- Wikforss, Åsa (2017). "Incomplete understanding of concepts." *Oxford Handbooks Online: Scholarly Research Reviews*.
- Williamson, Timothy (2007). *The Philosophy of Philosophy*. Oxford: Wiley-Blackwell.