

# The Myth of Social Content

Kirk Ludwig

## I

Social externalism is the view that the contents of a person's propositional attitudes are logically determined at least in part by her linguistic community's standards for the use of her words. If social externalism is correct, its importance can hardly be overemphasized. The traditional Cartesian view of psychological states as essentially first personal and non-relational in character, which has shaped much theorizing about the nature of psychological explanation, would be shown to be deeply flawed. I argue in this paper that social externalism faces insuperable difficulties. The first difficulty is that if syntax and semantics are independent, then social externalism is committed to the absurd consequence that many people have beliefs with formally inconsistent contents. I argue none of the possible responses to this objection are plausible. The second difficulty is that if syntax and semantics are independent, then social externalists are committed to a contradiction. After raising and defending these objections, I identify what I think is the underlying flaw in the social externalist position, a failure to pay attention to the structure of communication intentions, and briefly indicate how to neutralize the linguistic evidence social externalists advance in favor of their view.

Versions of social externalism are advanced in the work of Hilary Putnam<sup>1</sup> and, famously, Tyler Burge, in a series of influential articles beginning with "Individualism and the Mental."<sup>2</sup> I will choose Burge's more fully developed argument and position as my stalking-horse. I begin with a brief review

of one of Burge's well-known thought experiments in order to bring out that an essential assumption of the argument is that we can individuate an individual's attitudes by which sentences he would use to sincerely express them. It is only by means of this assumption that it could be shown that facts about the community's standards for the use of words could be shown to be relevant to the contents of a speaker's attitudes. This assumption will play a crucial role in the objections that follow.

## II

Burge's argument takes the form of a thought experiment in three steps. I quote in full the first and most important step.

A given person has a large number of attitudes commonly attributed with content clauses containing 'arthritis' in oblique occurrence.<sup>3</sup> For example, he thinks (correctly) that he has had arthritis for years, that his arthritis in his wrists and fingers is more painful than his arthritis in his ankles, that it is better to have arthritis than cancer of the liver, that stiffening joints is a symptom of arthritis, that certain sorts of aches are characteristic of arthritis, that there are various kinds of arthritis, and so forth. In short, he has a wide range of such attitudes. In addition to these unsurprising attitudes, he thinks falsely that he has developed arthritis in the thigh.

Generally competent in English, rational and intelligent, the patient reports to his doctor his fear that this arthritis has now lodged in his thigh. The doctor replies by telling him that this cannot be so, since arthritis is specifically an inflammation of joints. Any dictionary could have told him the same. The patient is surprised, but relinquishes his view and goes on to ask what might be wrong with his thigh. (p. 77)

In the second step, we imagine counterfactually our patient's history proceeds from the skin inward, non-relationally and non-intentionally described, just as it does, but that in contrast to the actual

situation his linguistic community uses ‘arthritis’<sup>4</sup> to apply to rheumatoid ailments that occur outside the joints, including rheumatoid ailments in the thigh. In the third step, we are invited to judge that, in the counterfactual situation, the patient does not have any beliefs which are correctly characterized using our word ‘arthritis’. “We suppose that in the counterfactual case we cannot correctly ascribe any content clause containing an oblique occurrence of the term ‘arthritis’” (p. 79).

Since in the first step of the thought experiment we described the patient as having beliefs involving the notion arthritis, although in the counterfactual situation, in which the patient’s non-relational and non-intentional history is the same, we describe the patient as having *no* beliefs involving the notion arthritis, it follows that the patient’s propositional attitude contents are not fixed solely by his history non-relationally and non-intentionally described. Furthermore, since the only difference between the situations we have described is in the use in the patient’s community of the word ‘arthritis’, his thought contents depend on how others in his community use the word ‘arthritis’. The conclusion applies of course to words we do understand and use correctly as well as to those we don’t. For if the community’s standards had been different, though our usage had remained the same, our thoughts contents would have been different.

Burge claims that the argument has extremely wide application, indeed, that it applies to virtually every word of the language, with two qualifications. First, the speaker must intend to be using his words as others in his community do. Second, he must not misunderstand them too wildly. Neither of these qualifications will affect the criticisms below.

Although Burge does not explicitly say that the patient in his example uses the word ‘arthritis’ to express his beliefs, unless we suppose this neither the example nor the thought experiment makes any

sense. For if the patient failed to use the word ‘arthritis’, it would be utterly mysterious how the community’s use of that word could have any relevance to what the patient thinks, or how the patient’s intention to use words as his community does could have any relevance to what he thinks on this occasion. The importance of the patient’s using the word ‘arthritis’ to express his beliefs comes out in the need to *identify* or *pick out* those beliefs of his over which the community’s influence is displayed in the thought experiment. Those beliefs are the ones he expresses using the word ‘arthritis’.

His misunderstanding, and consequently misusing ‘arthritis’, is also essential, for it provides the possibility of showing that the content of his belief tracks the community’s standards rather than what he believes to be the community’s standards. If what the patient believed to be the community’s standards never came apart from the community’s standards, it would not be possible to show which of these was the determinant of the content of the beliefs he expresses with the word ‘arthritis’.

This shows that the first step of the thought experiment makes an assumption equivalent to the conclusion. For in the first step of the thought experiment, we describe the patient as believing he has arthritis in his thigh when he says, “I’ve got arthritis in my thigh,” by attributing to him a belief whose content is given by the words he uses to express it *interpreted according to community standards*. So the question whether the conclusion is correct is the question whether it is correct to describe the patient in the first step of the thought experiment as believing he has arthritis in his thigh, or having attitudes involving the notion arthritis on other occasions, when he uses the word ‘arthritis’ to express his thoughts. The principle of attribution relied upon here can be stated as follows:

[P] The contents of a speaker’s attitudes are given by the propositions expressed by the sentences

he would use to express them as interpreted according to his linguistic community's standards for their use provided that (a) the speaker intends to be using his words in conformity with the standards of his community and (b) the speaker does not misuse of the words in the sentences too egregiously.

### III

In this section, I will show that principle [P] leads to unacceptable consequences. First, we will be led to attribute implausibly to the patient not only contradictory beliefs, but a belief whose content is contradictory. Second, we will be led to attribute, impossibly, two contents to a single belief. The only plausible retreat will show that it is a mistake to attribute contents to speakers by interpreting their words according to community standards, when they misunderstand or incompletely understand them.

These difficulties are direct consequences of two facts. The first is that the sounds and inscriptions we use as the outward form of thought acquire their semantic properties by convention, and hence are identifiable independently of the semantic properties they have, and could have had other semantic properties. This is the independence of syntax from semantics. The second is that speaking is a form of intentional action, which involves decisions about what words are appropriate to express one's thoughts to one's audience. Consequently, a speaker can choose on different occasions to express the same belief in different ways which he takes to be semantically equivalent.

A consequence of the independence of syntax and semantics is that in describing the beliefs of others we do not have to use the words they would use to express those beliefs. This is most obvious

in the case of attributing beliefs to someone in a language he does not understand. Suppose that I observer Karl, a monolingual German speaker, looking out the window, saying “Es Regnet,” and retrieving his umbrella from a closet. “Karl,” I correctly observe in English, “believes that it is raining” - though that is not how Karl would (or could) himself express his belief. There is nothing suspect about my belief attribution, despite my not using Karl’s words. The point holds also for people with whom I share a language. In attributing a belief to you, I do not have to use the words you would use to express that belief, as long as the words I use capture correctly the content of your belief. If you describe an acquaintance of yours as milk-livered, I can report your belief by saying you think him cowardly. Similarly, when reporting my own beliefs, I can do so in various ways. If I speak two languages, I can report a belief that it is raining by saying both “It’s raining” and “Es Regnet.” Or I can express a belief in English in two different ways, by saying, e.g., “Jack is milk-livered” or “Jack is a coward.”

The general point these examples illustrate is that we keep track of what others believe, hope, fear, desire and so on using our sentences, and aim to choose sentences whose meanings as used on the occasion correctly capture the role of the attitudes we ascribe in the subject’s psychology. The *sounds* or *inscriptions* we use are not what enables us to correctly track the subject’s thought contents, but rather the semantic, or if this is not fine-grained enough, notional import, in the context, of what we say.

To anticipate, and so to disarm, one possible objection here, we can note that there is a tendency, when we are attributing beliefs to speakers with whom we share a language, to attribute beliefs using the words that they themselves might use in describing their beliefs. The explanation for

this is simply that in attributing beliefs to them, we want to do the best job we can of telling our audience what to expect, and that includes what to expect in the way of verbal behavior. The case of cross-language attribution shows, however, that this is not necessary to correctly attribute beliefs to someone, and so is a pragmatic, rather than semantic feature of our attribution practices.

The first difficulty for Burge's argument has been noticed in the literature,<sup>5</sup> though not in quite the form developed here. It depends on the possibility of the patient being able to express the notion of arthritis using words other than 'arthritis', a possibility opened up by the independence of syntax and semantics. Suppose that the patient understands all of the words in the expression 'rheumatoid ailment of the joints', and that this is synonymous with 'arthritis'.<sup>6</sup> In this case, the patient will be able to express the notion of arthritis without using the word 'arthritis'. And there will be occasions, since we don't have to use his words in characterizing his beliefs, on which it will be appropriate to attribute to him beliefs using 'arthritis' obliquely in content clauses when there is no suggestion that he fails to realize his beliefs involve the notion of arthritis. In addition, it follows from [P] that all of the patient's beliefs that he would express sincerely using the word 'arthritis' also involve the notion of arthritis, despite his incorrectly or incompletely understanding that word.

Now consider what adherence to principle [P] has committed us to in the thought experiment modified as above. First observe that if we were to ask the patient, "Do you have a rheumatoid ailment of the joints in the thigh?", he would answer "No." Thus, given the principle of attribution [P], we should attribute to him a belief with the content: I do not have arthritis in my thigh. Here we use *our* words, rather than his, but this, as we have observed, is immaterial. The same content is attributed. Yet, if we were to ask the patient, "Do you have arthritis in the thigh?", he would say "Yes." Given

principle [P], we must attribute to him a belief with the content: I have arthritis in my thigh. Thus we have attributed to him contradictory beliefs. Having asked the subject these two questions, we could follow up by asking him our two questions at once: Do you have arthritis in your thigh and not a rheumatoid ailment of the joints? His answer, given what we have said about his metalinguistic beliefs, will be “Yes.” Thus, principle [P] commits us attributing to him a belief with the content: I have arthritis in my thigh and do not have arthritis in my thigh. In so rephrasing it, we reveal the structure of his beliefs at the level of notions or concepts. Thus we find we have attributed to him a belief with a contradictory content.

A view that threatens to regularly lead us to attribute contradictory beliefs to speakers must be implausible on the face of it, and one that threatens to regularly lead us to attribute to speakers beliefs with contradictory contents cannot be correct. Thus, we should reject the assumption that the patient who misuses ‘arthritis’ expresses with that word beliefs involving the notion arthritis.

We can consider briefly three responses to this objection: retreat, biting the bullet, and denying the legitimacy of the substitutions.

(1) Retreat. We can first try circumscribing the scope of the thought experiment by limiting it to cases in which the word a speaker uses but misunderstands expresses a notion he can’t express using other words. We can waive the objection that this response is *ad hoc*. The trouble is that it doesn’t retreat far enough. It is not enough to avoid the difficulty that someone should presently lack other words to express a notion *N* expressed by a word *W* he incorrectly understands. He must be unable to acquire any (other) words to express *N*, that is, he must be unable to acquire the notion *N*. For if he does acquire other words to express *N*, without correcting his misunderstanding of *W*, we will be led

again, first, to attribute to him contradictory beliefs, and, second, to attribute to him a belief with a contradictory content. Thus, retreat would require that the speaker have a belief involving the notion  $N$  when he expresses a belief using  $W$  only when *he is incapable of having his misuse corrected*. We have now clearly left behind the first step of the thought experiment; and it is distinctly odd to suggest that it is appropriate to attribute to a speaker a belief involving a notion which he is *unable to acquire*.

(2) Biting the bullet. Second, we can agree that most people do have (or can be led to have) contradictory beliefs, and beliefs whose contents are contradictory, precisely because they misunderstand many words they routinely use. In response, we first can observe that this is not made more plausible by the possibility of having pairs of beliefs which are cannot logically be jointly true, or the possibility of having necessarily false beliefs, as in the case of false beliefs about mathematical equivalences. In attributing such beliefs to someone, we are not licensed to represent the form of the beliefs as either  $X$  believes that  $p$  and  $X$  believe that not- $p$ , or  $X$  believes that  $p$  and not- $p$ . So the intelligibility of these kinds of beliefs does not help make intelligible the beliefs we would be led to attribute following principle [P]. Secondly, since a rational agent will not knowingly believe something that is self-contradictory, we must hold the speaker is either irrational, or doesn't know what he believes. There's nothing (necessarily) irrational in misunderstanding a word. So we must conclude that the speaker doesn't realize what he believes (not just what he's saying, which is a different matter).<sup>7</sup> In my opinion, this does far more damage to our ordinary conception of belief content than would rejecting the first step of the thought experiment.

(3) The third response is to reject the principle of the intersubstitution of synonyms in content clauses. Burge, in particular, has argued against this principle in "Belief and Synonymy"<sup>8</sup> and so we

should see whether denying this principle really does avoid the problem. It does not, because the only reason to deny that synonyms can be substituted in content clauses would be the claim that synonyms don't always express the same notions.<sup>9</sup> As long as different expressions can express the same notions, whether or not all synonymous expressions do, the same difficulty can be raised. So denying this does not avoid the problem. One would have to deny that different expressions can express the same notions, which is to deny the independence of syntactical and semantic properties. The reasons for adhering to this principle have already been given.

The second difficulty relies on the possibility of a speaker expressing one of her beliefs in more than one way. One will be disposed, other things being equal, to express a belief using different expressions just in case one believes the different expressions to be synonymous, or, to put it more cautiously, just in case one believes the different expressions express the same notions (or contents). There are two cases to consider, first, that of someone who treats others of his expressions which he understands as synonymous with (or expressing the same notion as) a word he misunderstands, and, second, the case of someone who isn't prepared to treat any of his other words or expressions as synonymous with the word he misunderstands.

(1) In the first case, he is disposed to express the same belief using *W* and using the expression (*W\**) which he thinks synonymous with *W*. *W* and *W\** are not synonymous according to his community's standards. Therefore, if we assign content to his beliefs by assigning to them the content appropriate for the words he uses to express them, interpreting his words according to community standards (provided that he intends to be using his words as others in his community do, and does not grossly misuse them), we will attribute to one belief two different contents. Since we individuate beliefs

by their contents, it is not possible for one belief to have two different contents, so this leads directly to a contradiction.

Since we must give up one of these attributions as incorrect, it seems clear that we should attribute the content corresponding to the speaker's expression of his belief in words he fully understands. If we corrected the speaker's mistake about *W*'s proper use, he would stop using *W* to express that belief, but continue to use *W*\*. The content of his belief does not change, although we now correctly attribute it using *W*\*. So its content all along was correctly attributed using *W*\*. The description of the patient in the thought experiment's first step will therefore be incorrect if the patient thinks any expressions he fully understands are synonymous with 'arthritis'.

(2) It might still be thought, however, that the thought experiment can get going in cases in which the patient does not think any of his words or expressions are synonymous with 'arthritis'. This is our second case. But this does not avoid the difficulty. If a speaker who misunderstands a word does not now have any other expressions to express what it means in his community, he can acquire them, in which case the same difficulty arises. We must decide which expression correctly captures the content of the belief he expresses with either indifferently, and we will choose in favor of the one he fully understands. It does not matter that a speaker does not yet have such an expression, for his acquiring it does not change the contents of his beliefs, but only gives him new expressive resources. After he acquires a new expression he fully understands and thinks synonymous with the word he misunderstands, we would attribute a content to his belief corresponding to the acquired expression. Since the content of his belief would not have changed, prior to his acquiring the new expression the content was incorrectly ascribed using the word he misunderstood.

It is therefore never appropriate to attribute to a speaker a content by interpreting according to community standards the words he uses to express his belief, when he does not fully or correctly understand those words, and, consequently, principle [P], upon which the argument for social externalism must be based, is false.

#### IV

When we focus on speaking in the context of communication, we can see why it is a mistake to read off from what words a person uses to express a belief its content by interpreting them, even within the constraints given in [P], according to the community's standards. The basic difficulty lies in the fact that we can distinguish between the content of a belief (or other propositional attitude) and the words used to express or describe it. This makes possible our misexpressing a belief because what we thought a word meant to our *audience* is not what it did. In talking to other people, our choice of words is guided by what we intend them to come to believe we are asserting, and by how we think they will understand the words we assert. Mistakes occur when we are wrong about how they will understand the words we use. But this mistake could not convict us of believing on that occasion what those words expressed in our community or to our audience. For our mistaken choice of words was guided by *antecedent knowledge of the belief we wished to express*, and our mistake was to suppose the words we chose correctly expressed that belief. That speaking is a form of rational action which requires deliberation about how an audience will understand verbal behavior explains why knowledge of contents is a regulative ideal in interpretation, and why it is a mistake to think someone

believes what he literally says when he doesn't understand his words as you or others do.

If this is right, then how should we understand what we typically say when corrected in our improper use of expressions? Consider the following possible conversation:

Patient: Doctor, I think my arthritis has spread to my thigh.  
Doctor: That's not possible. Arthritis is a rheumatoid ailment of the joints. That it occurs only in the joints is a part of its meaning [sic].  
Patient: Oh, I didn't know that! I guess I was mistaken to think that I had arthritis in my thigh. Well, anyway, what should we do about my thigh? I can hardly walk.  
Doctor: Just put your prescription arthritis balm on it. That should help relieve the inflammation. Call me next month to let me know how it's going. You can pick up your bill on the way out.

The difficulty is that the patient seems to be saying that he had the false belief that he had arthritis in his thigh. This, if we take him to be correctly characterizing his belief, is in conflict with our conclusion that his beliefs were not correctly described using 'arthritis' as interpreted according to community standards. The correct response is to observe that semantic ascent is hard. The patient's verbal response is confused in a way akin to the doctor's use/mention confusion. The patient does not explicitly signal in what he says that he is concerned with the word, simply because he, along with most people, are not used to shifting from talk about things to talk about the expressions they use to talk about things. So a mistake in the use of an expression is apt to get expressed as a mistake in a belief about things. That this is the right interpretation is shown by what the patient would say if pressed about whether he *really* thought he had a rheumatoid ailment of the *joints* (i.e. arthritis) in the thigh. I think it is clear he would deny having thought that. This will force us to interpret what he says as an admission of error about what the word 'arthritis' means, rather than an admission of error in thinking

he had a rheumatoid ailment of the joints in the thigh.

*Department of Philosophy  
University of Florida  
Gainesville, FL 32611  
USA  
352-392-2084  
kludwig@phil.ufl.edu*

## Notes

1. See "The Meaning of "Meaning"" in *Mind, Language and Reality, Philosophical Papers*, vol. 2, (Cambridge: Cambridge University Press, 1975), pp. 215-71.
2. In *Midwest Studies in Philosophy*, IV, edited by French, Uehling, and Wettstein (Minneapolis: University of Minnesota Press, 1979), pp. 73-121. All parenthetical citations to page numbers are to this article.
3. A word or expression occurs obliquely if existential generalization fails if it is a singular term, or substitution of coreferring terms or coextensive terms may fail to preserve the truth value of the sentence containing the term.
4. I use single quotation marks throughout to form quotation names of expressions and double quotation marks for direct quotation, reports of speech, and article titles.
5. Jerry Fodor, "Cognitive Science and the Twin Earth Problem," *Notre Dame Journal of Formal Logic*, 23, July 1982, pp. 98-118. Burge replies to Fodor's criticism in "Two Thought Experiments Reviewed," op. cit. pp. 284-293. The form in which I put the objection avoids the objections Burge raises to Fodor's criticism.
6. We can treat these expressions as synonymous for the purposes of the paper. Nothing will turn on this being exactly right, as long as it is at least possible to introduce into the language a synonym for 'arthritis'.
7. This is a particularly vivid way of bringing out why social externalism of Burge's variety poses a problem for self-knowledge. These issues deserve fuller discussion, however, than I can give them here. It can be clearly shown, I believe, that Burge's defense against this charge in "Individualism and Self Knowledge," (*Journal of Philosophy*, 85, November 1988, pp. 649-663) is inadequate, and that the kind of error involved cannot be assimilated, as Burge would suggest, simply to the inability to explicate our concepts in other words. That is not what is going on here.
8. *Journal of Philosophy*, 75, March 1978, pp. 119-138.
9. One ground for this worry might be that although articulated and unarticulated expressions may be synonymous, the use of an articulated expression may indicate a corresponding articulation of the belief content. The argument, however, does not depend on picking an articulated expression as the synonymous expression the patient understands.