

[Draft, October 2004, do not quote or distribute without permission.]

THE BURGEAN INTUITIONS

David Pitt

California State University, Los Angeles

The thought experiments advanced by Tyler Burge (1979, 1982, 1986) are widely held to have established that psychological content can be socially determined.¹ They are taken to show that, contrary to the traditional Cartesian conception, the contents of an individual's thoughts are not always determined by his intrinsic properties alone, but can depend on the practices of the linguistic community of which he is a member. Specifically, Burge claims that his thought experiments show that the socially determined meanings of the words a linguistically competent individual uses to express his thoughts about himself, his fellows and his physical environment can determine the contents of those thoughts, and, hence, that those contents themselves can be socially determined.

In this paper I argue that Burge's anti-individualist thesis is not supported by his thought experiments. The cases Burge presents are meant to elicit intuitions which, in combination with a plausible general principle about belief ascription, provide a strong motivation for abandoning the individualist conception of mind. The intuitions concern *what it is natural to say* about what the individuals described in the thought experiments believe, and the general principle is that *all things equal* belief-ascriptions it is natural to make are literally true. The intuitions are supposed to override any sense that the *ceteris paribus* clause of the principle is sprung in Burge's cases because of the conceptual idiosyncracies of his subjects: they believe what they say *in spite of* their deviance from the communal norms governing the usage of their words.

The intuitions about what it is natural to say are undeniably powerful. Burge is surely right that in the situations he describes it is natural to attribute beliefs to his subjects using the very words they utter, even though their understanding of those words is at odds with their socially determined meanings. Nor would the subjects themselves hesitate to describe their thoughts in the same way. Nonetheless, I maintain, these intuitions are not sufficient to establish Burge's thesis. What it is natural to *say* in a given case need not be literally *true*; and there may be compelling reasons for thinking that it is not – even if it remains quite natural to say it. I will argue that in any case of the kind required by a Burgean thought experiment there is a reason for not taking the homophonic ascriptions at face value, which is not neutralized by the intuitive naturalness of making them. Given that it is the strength of these intuitions that accounts for the persuasiveness of the thought experiments, if it can be shown that they do not in fact ensure the literal truth of the ascriptions, Burge's case for linguistic anti-individualism ought to seem much less compelling.

I. THE STRUCTURE OF BURGE'S ARGUMENT AND THE ROLE OF THE NATURALNESS INTUITIONS.

The main argument of "Individualism and the Mental" is as follows.² Bert (as I'll call him) is a member of our linguistic community who has arthritis. Bert speaks of his ailment using the term 'arthritis', which in English applies by definition to diseases of the joints only.³ He says things like "I have arthritis," "The arthritis in my wrists is more painful than the arthritis in my ankles," "My grandmother had arthritis too," and so on, and thereby expresses his (true) beliefs about his condition. One day he develops a condition in his thigh that he thinks is the same as the one he has in his wrists and ankles. He visits his doctor and says "I think my arthritis has

spread to my thigh.” The doctor informs him that this could not be true, since arthritis cannot occur in muscle. Bert stands corrected, and is relieved that he was wrong to think he has arthritis in his thigh.

Next we are to imagine Bert in a counterfactual situation in which his physical history is unchanged up until the moment he expresses his worry to his doctor, but in which he is a member of a different linguistic community, in which ‘arthritis’ applies by definition to rheumatoid diseases occurring in muscle as well as joints. In this counterfactual context, none of Bert’s ‘arthritis’ utterances mean what they do in his actual context, and none of his beliefs have their actual content. Since by hypothesis the only difference between the actual and counterfactual situations resides in the practices of the linguistic communities Bert is a member of, we seem compelled to conclude that it is these factors, and not something internal to Bert, that determine the contents of his ‘arthritis’ beliefs (the beliefs he would express using the term ‘arthritis’). More specifically, in both cases the socially determined meanings of the words Bert uses in reporting his beliefs determine his beliefs’ contents. Hence, those contents themselves are socially determined. Moreover, since similar cases can be described for a very large variety of concepts, the conclusion generalizes: mental content is in general determined socially. To summarize:

1. In the actual context, Bert believes that he has arthritis in his thigh.
2. In the counterfactual context, Bert does not believe that he has arthritis in his thigh.
3. Bert’s physical constitution is the same in the actual and counterfactual contexts.

Hence,

4. The difference in the content of Bert’s thoughts is not determined by any of his

intrinsic
physical features.

But,

5. The only relevant differences between the actual and counterfactual contexts concern the practices of the respective linguistic communities.

Hence,

6. Bert's social/linguistic context determines the contents of his thoughts.

More specifically,

7. The socially determined meanings of the words Bert uses to report his thoughts determine the contents of those thoughts.⁴

It is thus essential to Burge's argument that premise (1) be literally true. If Bert doesn't believe this in the actual case, then the counterfactual case is irrelevant, and no conclusions about content determination are established by the thought experiment. (In the counterfactual case there is no discrepancy between what 'arthritis' means and what Bert thinks it means.) It is only on the assumption that Bert believes he has arthritis in his thigh that the content-determining role of social context is made salient by a comparison with the counterfactual case. In spite of the fact that Bert's intrinsic physical features are held constant, he does not say or think in the counterfactual case what he says and thinks in the actual case. Since the only relevant difference between the contexts concerns the extrinsic (to Bert) facts about what 'arthritis' means in the two linguistic communities, it would seem that it must be these social-linguistic facts that determine the contents of Bert's thoughts.

However (as Burge notes), one might think that the very fact that Bert says "I think I have arthritis in my *thigh*" is evidence that he *does not* have the concept ARTHRITIS – since it is a

conceptual/definitional truth (in some non-committal sense) that arthritis is a disease of the joints only – and, hence, that he cannot have the belief that he has arthritis in his thigh. After all, in the counterfactual case Bert does not have the concept ARTHRITIS *because* ‘arthritis’ in the counterfactual community applies by definition to a disease that can occur in muscle as well as joints. So why shouldn’t his belief that ‘arthritis’ applies by definition to (we may suppose) those very muscle ailments preclude him from having the concept? It might seem (as the traditional view of concept-possession has it) that in order to have a concept *C* an individual must believe (or be disposed to believe) that *Cs* are *F* (and other related propositions), for all *Fs* that are part of or are entailed by the definition of *C*, and for no *Fs* such that not-*F* is entailed by the definition of *C*. Hence, no one who uttered ‘arthritis is *F*’ for any *F* precluded by the definition of ‘arthritis’ could thereby be expressing the concept (thought-constituent) ARTHRITIS, since any such utterance would betray a lack of understanding of the meaning of ‘arthritis’. So why does Burge think we should nonetheless allow that Bert has the concept?

The answer is that we allow that Bert has beliefs of which ARTHRITIS is a constituent. He uses the word ‘arthritis’ to say true things like “I have arthritis in my ankles,” “Arthritis is a disease,” “Arthritis is painful,” etc. He is not, to our knowledge (before his visit to the doctor), an abnormal speaker of English; he is not obviously deceitful, dissembling or deranged; there is nothing odd about his other utterances; etc. In these cases we take Bert’s utterances at face value: what he says (the contents of his utterances) is what he believes (the contents of his thought). When we say that he has these beliefs, we take our own ascriptions at face value as well: what he believes (the contents of his thoughts) is what we say he believes (the contents of our utterances). We have no reason not to. As Burge says: “there is a methodological bias in

favor of taking natural discourse literally, other things being equal. ... Literal interpretation is *ceteris paribus* preferred.” (88) But Bert’s utterance of “I think I have arthritis in my thigh” is *prima facie* problematic. It betrays a misunderstanding of ‘arthritis’, which by definition cannot occur in muscle tissue. Bert has not properly understood all of the words he has used. Shouldn’t this provide reason *not* to take what he says, and what we say of him, literally? Why doesn’t his confusion spring the *ceteris paribus* clause in the principle of literal interpretation? Why don’t we say that all things are not equal because Bert doesn’t understand what he has said?

The reason we don’t back away from literal interpretation in this case – the reason that all things remain equal in spite of Bert’s mistake – is, according to Burge, that there is a very strong intuition that the correct way to describe Bert’s mistake is to say that he believes that he has arthritis in his thigh. It is entirely natural for us to describe Bert – and for him to describe himself, even after he is apprized of his error – as believing (or having believed) that he has arthritis in his thigh. This is simply what we say, and it is entirely natural. So if Bert says “I have arthritis in my thigh,” and it seems entirely appropriate and unproblematic for us to attribute to him a belief using the very words he uttered, then we ought to take those words at face value: Bert (mistakenly) believes that he has arthritis in his thigh. But since he cannot believe this unless he has the concept ARTHRITIS, he must have that concept, in spite of his misunderstanding.

Now, a moment’s reflection on these considerations shows that in fact the contrasting counterfactual case is an inessential component of Burge’s argument. The reasons for thinking that the ascription to Bert of the belief that he has arthritis in his thigh is literally true, despite his confusion, are themselves reasons for thinking that psychological content is socially determined.

Given that literal interpretation is preferred even in cases of partial or misunderstanding such as Bert's, and that literal interpretation is determined socially, it follows that the contents of Bert's thoughts are determined socially.⁵ The comparison with the counterfactual case is simply a way of dramatizing this conclusion by providing a case in which a *different* content is determined by a *different* linguistic context. The heart of the argument is the explanation of *how* content is determined by social/linguistic context. Indeed, Burge seems to recognize the general point in Burge 1978 (134-35):

Our discussion has brought out the domineering role of the presupposition that speakers are to be taken at their word; that is, literally or homophonically. ... When the presupposition is in force, communal conventions about the meaning of a speaker's words tend to override what a speaker mistakenly associates with his words in determining what he says and even, sometimes, what he believes.

Burge's case for social anti-individualism is just the case for the literal truth of the ascription to Bert. It rests upon the intuitive naturalness of the way we describe Bert, and the principle that, all things equal, literal interpretation of natural discourse is preferred (and not, as it might appear, on the comparison of the actual and counterfactual cases).⁶

A further point emerges here. The mechanism whereby mental content is determined socially in Bert's case *in spite of his misunderstanding* ought to work in perfectly ordinary cases as well.⁷ In the normal course (i.e., all things being equal), we take at face value propositional attitude ascriptions we make on the basis of what we believe to be sincere utterances. We take people at their word when they say what they think: the content of their thoughts, the linguistic content of their utterances, and the linguistic content of our ascriptions match. What we say is what they say, and what they say is what they think; hence, what we say is what they think. Moreover, since the contents of our utterances – their linguistic meaning – are determined

socially, the contents of their thoughts are determined socially as well.

Consider a case in which Bert has complete and accurate knowledge of the communal meaning of 'snow'. He has said a lot of true things about snow. He says "Snow is cold," "Snow is white," "Snow falls from the sky in winter," "Snow is frozen rain," "Snow is fun to play in," etc., and no false or odd things about it. So we have good reason to think that Bert is competent with the word 'snow'. Now focus on a particular (sincere) utterance of Bert's, say "Snow is cold." On the basis of this utterance, we (naturally and compellingly) attribute to him the belief that snow is cold. That is, we take the content of our ascription to be the content of his thought. What did Bert think? He thought that snow is cold. How do we know that is what he thought? Because he said "Snow is cold," 'snow is cold' means that snow is cold, this is the form of words it is natural to use in ascribing the thought to Bert, our ascription has the content that Bert thinks that snow is cold, and our ascription is true. Therefore, Bert thinks that snow is cold. The content of his thought can be *read off* the content of his utterance and our attribution. What he says is what he thinks. However, since the content of what he says, and thus the content of our homophonic attribution, is determined socially, the content of Bert's thought is determined socially.

But this argument is unconvincing. It doesn't show that the match between utterance content and thought content is not simply due to the fact that, normally, competent speakers choose words that correctly express their thoughts, and that *ceteris paribus* one should assume that they do. Contrary to the argument's conclusion, there is nothing in the situation as described to suggest that linguistic content *determines* mental content. There is a simpler and more plausible explanation for the identity of utterance-content and thought-content.

If, however, it can be shown that the socially determined linguistic content of a thinker's utterance (and of the third-person ascription) can override misconceptions on the part of the thinker, in the sense that the thinker's mistaken thoughts about the meanings of the words he uses lose out to the actual meanings of those words in identifying the thoughts he expresses with them, then we would have a potentially compelling reason to think that linguistic content determines mental content. And this seems to be just what Burge's cases are designed to do. They are supposed to provide the contrast needed to bring out the determinative role of a speaker's language with respect to his thought contents: *even when a speaker misunderstands* the words he uses in expressing his thoughts, the socially determined meanings of those words is still the content of his thoughts. In these cases the explanation that the speaker has merely chosen the correct words to express his thought is not available. For it assumes that the speaker knows the meanings of those words, whereas by hypothesis Burge's subjects do not. Hence, there is some reason to think that just as one can *say* something one does not understand, one can *think* something one does not understand, and that the explanation for how this is possible is the same in both cases: what one says and thinks is determined by the communal meanings of one's words.

So, the deeper structure of the argument in "Individualism and the Mental" involves the presentation of cases which by virtue of their dramatic contrast with ordinary ones bring out the constitutive role of our ordinary belief-ascribing practices in the determination of mental content in general. The cases, and the argument, are not essentially different from the example Burge presents in "Belief and Synonymy" of his own misunderstanding of 'fortnight'. Here too (as is clear from the passage quoted above) we have a contrast case that brings out the determinative

role of communal meaning. (Having once mistakenly thought that ‘fortnight’ is defined as “a period of 10 days,” Burge describes himself as having believed that a fortnight is a period of 10 days. The description seems at once entirely natural and literally true.)

These cases are also, despite obvious differences, in essence the same as the one Burge constructs in “Intellectual Norms and Foundations of Mind.” There we have a subject (call him “Al”) who is a victim of neither partial nor misunderstanding with respect to the specimen term (‘sofa’). Al understands the meaning of the word ‘sofa’ in his language perfectly well. He knows that sofas are in his linguistic community taken to be by definition pieces of furniture, and he is fully competent in the use of the word. Nonetheless (according to Burge), he doubts that sofas are furniture. Al has become convinced that they are in fact some sort of works of art or religious artifacts, not meant for the use some people put them to at all, and he expresses his suspicion by saying “Sofas are not pieces of furniture!” He is wrong, however. His belief that sofas are not furniture, which was formed on the basis of systematically misleading evidence, is false. In a counterfactual case, however, we are to suppose that a physically identical Al says something *true* when he utters the sentence ‘sofas are not pieces of furniture’. In the counterfactual language community, the word ‘sofa’ is applied by definition not to pieces of furniture, but to works of art or religious artifacts, just as Al suspects. (He’s wrong in thinking he’s made a startling discovery, however, since everyone already knows this.) When Al counterfactually says “Sofas are not pieces of furniture,” his utterance, and the belief it expresses, are both true. They do not, however, have the same content as Al’s actual utterances. He counterfactually neither says nor believes that sofas are not pieces of furniture, because ‘sofa’ in the counterfactual community doesn’t mean *sofa*.

The pattern of reasoning in Al's case is the same as in Bert's: the difference between the contents of Al's actual and counterfactual beliefs is determined by the actual and counterfactual meanings of the words he utters, and those meanings are determined by factors extrinsic to Al. Hence, Al's thought contents are determined by factors extrinsic to him. Moreover, the points made above with respect to the role of the intuition about what Bert believes in the actual case, and the comparison to the counterfactual case, apply to Al's case as well. The considerations that are supposed to lead us to think that Al believes that sofas are not pieces of furniture, *in spite of the fact that he knows that the word 'sofa' in his community applies by definition to pieces of furniture*, are themselves sufficient to show that the content of Al's sofa thoughts is socially determined. The comparison of actual Al and counterfactual Al is not what establishes the anti-individualist conclusion.⁸ It merely dramatizes it.

Further, the point of the case of Al seems to be to provide contrast with ordinary cases as well, since the principles Burge appeals to ought to establish anti-individualism in *any* case in which natural discourse is taken at face value. However, in Al's case it is not misunderstanding that is overridden by communal meaning, and which provides the essential contrast to the ordinary case – since, by hypothesis, Al understands perfectly well what his community takes 'sofa' to mean. It is, rather, as Burge says, "nonstandard theory" (1986: 709). Nonetheless, there is an important parallel to Bert's case. Both Bert's and Al's utterances are *prima facie conceptually* problematic: neither involves a merely empirical divergence from communal standards and belief. Bert's utterance indicates conceptual confusion (likewise Burge's utterance of 'a fortnight is a period of 10 days'), and Al's indicates an attempt at conceptual subversion. I will characterize what the cases have in common as *conceptual dissonance*. In all of Burge's

examples it is the fact that the individuals in question think what they say *in spite of the conceptual dissonance of their utterances* that provides the reason for thinking that the principle of literal interpretation has general application because linguistic meaning determines mental content. The fact that linguistic meaning trumps individualistic factors in cases of dissonance between what a speaker believes about his words and what they mean in his language shows why there is normally a match between linguistic content and mental content: it is not because mental content determines linguistic meaning, or because speakers normally choose their words correctly. The contrast between ordinary and dissonant cases brings out the “domineering role” of the principle of homophonic interpretation.

Now, the inclusion of the *ceteris paribus* clause indicates that Burge thinks there are conditions under which literal interpretation *is not* preferred – cases in which the principle does not override the conceptual dissonance of the thinker’s utterance. In such cases, the individual is not to be taken at his words. This might mean that the individuals’ words are themselves to be reinterpreted: perhaps one should suppose they are not after all members of one’s linguistic community; perhaps they are speaking a foreign language, or a different dialect of English. If, however, one is confident that a speaker is a linguistic compatriot, it is more likely that his words will be taken to mean what they mean in the language he speaks, but that they will not be taken to be an accurate expression of what he *thinks*. In such cases, the link between the meaning of the words one has uttered and the content of the thought one wished to express is broken or otherwise disabled.

Indeed, if we were to hold that there *cannot* be a mismatch between linguistic meaning and thought content, then we would be contradicting some very obvious facts. For what then

would we say of malaprops, slips of the tongue, phonemic mispronunciations, etc.? To deny that such phenomena justify reinterpretation would seem to imply a kind of infallibility (enforced by language) with respect to one's choice of words in expressing one's thoughts. One simply could not misexpress them. If one's utterance is sincere, then one must be thinking what one's words mean. But this is clearly false; and any view that is committed to it ought to lack any intuitive appeal.⁹

If there can be cases of mismatch between thought content and linguistic content, then, it remains to be seen what sorts of criteria there are for identifying them. What kinds of considerations override our default commitment to taking our fellow speakers at their words? Burge (1978: 134; 1979: 90-1) mentions utterances involving slips of the tongue, spoonerisms, malaprops and radical misunderstandings as appropriate candidates for reinterpretation. Presumably, at least some of these involve a malfunction of speech-processing mechanisms, and may be to that extent involuntary. Moreover, speakers typically correct themselves ("That's not what I meant!") when they realize or are told of their mistakes. So in such cases there seems to be sufficient reason to excuse them from their *prima facie* commitment to having thought (or meant) what they said. It would be uncharitable to hold a speaker to his words in such cases.

In a parenthetical remark (1979: 91), however, Burge indicates that in fact he thinks utterances involving malaprops or radical misunderstandings¹⁰ are *not* exempt from the principle of literal interpretation. He says that he is "not convinced" that someone who believes that 'orangutan' is a word for a fruit drink and says "An orangutan is a fruit drink" should not therefore be taken to *mean that* an orangutan is a fruit drink and *think that* an orangutan is a fruit drink. If Burge's argument for anti-individualism really is committed to this, then I think it can

simply be dismissed. If a case of misunderstanding this radical doesn't trip the *ceteris paribus* clause, then it's hard to imagine what would. Burge claims that his thought experiment depends upon completely ordinary, completely compelling intuitions. But if it depends upon a principle that commits us to the claim that the communal meaning of *any* sincere well-formed utterance is the content of the thought it expresses, then it is completely *counterintuitive*.

But it seems that Burge could reject such cases and still have an important thesis about the determination of mental content.¹¹ He says (1979: 92): "The thought experiment depends only on there being some cases in which a person's incomplete understanding does not force reinterpretation of his expressions in describing his mental contents. Such cases appear to be legion." (92) I take it Burge means that his argument does not require commitment to a principle of literal interpretation of every utterance. Standard practice can call for reinterpretation in cases of linguistic incompetence or momentary lapses without spoiling the anti-individualist argument. In such cases all things are not equal, and our commitment to taking speakers at their words is suspended. What the argument requires is cases in which the principle of literal interpretation overrides conceptual dissonance of one (not-too-serious) kind or another. Conceptual dissonance is a *prima facie* reason for thinking that all things are not equal. This is why Burge stresses that the deviance from community norms is not merely empirical. Mere empirical deviance would not raise the issue of non-literal interpretation, and would therefore not provide the resistance needed to reveal the determinative role of the principle. Since homophonic interpretation is still preferable in these cases, they do show something interesting and important about psychological content.

But *why* should we think that the principle of literal interpretation (henceforth, the

“Principle”) applies in spite of the *prima facie* evidence that things are not all equal for Burge’s subjects? Why isn’t its *ceteris paribus* clause tripped by the dissonance of their utterances? How is the resistance to homophonic interpretation overcome? The answer constitutes the very heart of Burge’s argument for anti-individualism. We do not reinterpret in these cases – we do not choose different words to match the subjects’ deviant conceptions – because of very powerful intuitions about the correct way to characterize them. Bert says “I’m afraid I might have arthritis in my thigh.” We know that what he fears cannot be true; we know he has made a mistake. Yet we describe his mistaken belief using the very words he doesn’t understand: we say that he fears he has arthritis in his thigh. Burge says “I used to believe that a fortnight is a period of ten days.” He knows what he used to believe cannot be true; he knows he made a mistake. But he describes what he used to think using the very word he didn’t understand at the time – he says that he used to think that a fortnight is a period of ten days. Al thinks everyone is wrong about what sofas are. He knows what ‘sofa’ means in his language. He knows that it’s (something like) definitional that sofas are pieces of furniture. Yet he says “Sofas are not pieces of furniture,” and he and we interpret his utterance literally. We might think he is wrong, but we describe him as thinking that sofas are not furniture. These are, as Burge stresses, the utterly natural and intuitively correct things to say. No one without a theoretical ax to grind would question their appropriateness. But this shows that communal meaning trumps individual misconception in determining the contents of a generally competent speaker’s thoughts. It’s a philosophically surprising but inevitable consequence of our ordinary practice of ascribing thoughts on the basis of sincere utterances that mental content is determined by factors extrinsic to thinkers.

The role of the intuitions Burge evokes (henceforth, the “Intuitions”) should now be clear.

The Principle applies in these cases in spite of *prima facie* evidence that all things are not equal because of the powerful intuition that our homophonic ascriptions are correct. The intuitive naturalness of our ascriptions to Bert, Burge and Al prevents the *ceteris paribus* clause of the principle from being sprung. In spite of the fact that their utterances are conceptually dissonant, it still seems perfectly appropriate to use their words in characterizing what they think. The power of Burge's thought experiments lies in the strength of these unspoiled intuitions.

The Intuitions are undeniably powerful. I share them. It is overwhelmingly natural to characterize mistakes like those made by Bert, Burge and Al using their very words. My own informal survey has shown that Burge is right about the philosophically untutored. I also find a very strong tendency for homophonic description in my own case. Nevertheless, I think the Intuitions are misleading. They cannot be uncritically accepted as evidence that all things are equal in the situations Burge describes. Countervailing considerations show that what it is overwhelmingly natural to *say* in Burge's cases cannot, on reflection, be taken to be literally *true*. The Intuitions do not, after all, prevent the *ceteris paribus* clause of the Principle from being sprung. Some of these considerations have been discussed in the literature on Burge's thought experiments (though not always, in my view, as effectively or persistently as they ought to have been), and some are presented here for (to my knowledge) the first time.

Though Burge's precept of upholding natural intuitions in the face of theoretical assault should not be rigidly adhered to (after all, it did once seem intuitively obvious the sun moves around the earth, that the earth is flat and that the natural state of motion of physical objects is rest!),¹² I think the best way to challenge Burge is with intuitions as natural and powerful as the ones he offers. Since Burge's case for anti-individualism rests so fundamentally on the

commonsense intuitions he evokes, meeting it on its own terms requires equally natural commonsense intuitions and principles whose application in the situations he describes shows that his conclusions are unjustified. Of course there is no guarantee that a given commonsense intuition can be met with an equally plausible counterintuition. The intuition may simply be correct and indisputable, or refutable only by theory. I don't think this is the case with the Burgean Intuitions, however. In what follows I discuss three features of our ordinary belief-ascribing practices that militate against Burge's anti-individualist conclusion. These practices are governed by a principle of charity enjoining ascription of contradictory or incoherent beliefs except in extraordinary circumstances, and also by the principle that utterances betraying linguistic incompetence should not be interpreted literally. In addition, there are many cases in which homophonic ascription is quite natural, but in which there is also a very clear commonsense intuition that on (brief) reflection literal interpretation is not intended. Taken together, these three features make a strong case against Burge's argument that all things are equal in his cases with respect to the principle of literal interpretation.

II. FIRST OBJECTION: THE PRINCIPLE OF CHARITY.

It has been objected to Burge (beginning with Fodor 1982¹³) that all things are not equal in Bert's case because the sentence 'he has arthritis in his thigh' is, by hypothesis, self-contradictory or conceptually incoherent, and there is a competing principle governing ordinary discourse that says that one's attributions should be charitable. One should not, *ceteris paribus*, attribute contradictory or incoherent beliefs. But to attribute a belief to Bert using the sentence he uttered is to attribute to him belief in a contradiction. Perhaps he could have such a belief¹⁴; but the default assumption – the intuitively very natural assumption – is that if a speaker

assertively utters a sentence which in the language he speaks is contradictory or incoherent, then he has misexpressed what he thinks. Only in unusual circumstances would we be willing to suspend the presumption of rationality and attribute a contradictory or incoherent belief.

The question whether or not Bert has such a belief should be distinguished from the question whether or not we attribute one to him in using a homophonic ascription. If Bert utters the sentence ‘my arthritis has spread to my thigh’, it might be open to question what he actually meant (what he actually thought). But if *we* utter the sentence ‘Bert believes that he has arthritis in his thigh’, then there is no question that the embedded sentence ‘he has arthritis in his thigh’ is *for us* conceptually incoherent. For, as Burge makes clear, Bert’s error is not an ordinary empirical one; it is a *conceptual* matter that arthritis cannot occur outside of joints. Given that it is reasonable to suppose, as Burge does (1979: 102), that the kinds of errors he describes are common, to suppose that individuals’ words should be taken at face value in such cases is to turn a practice of intending to make sense of their behavior (belief attribution) into a practice that frequently makes nonsense of their mental lives. But if this is not our practice, then the kinds of cases Burge describes in “Individualism and the Mental” are ones in which literal interpretation is not preferable, and the conclusion he wants about the contents of thought does not follow from the naturalness of our attributions. The fact that we find it so natural to make them in spite of their incoherence no doubt requires an explanation.¹⁵ But if the principle of charity is upheld, it cannot be maintained that we ought to take them to be literally true.

In discussing another example, Burge (1982a: 290) responds to Fodor by claiming that Fodor mistakenly attributes to him the view that “the fact that contracts need not be written is constitutive of our concept” of contract, whereas he “assumed only that it is not ‘constitutive of

our concept' of contract that contracts must be written." The concept of a contract leaves it unspecified whether or not contracts can be unwritten. Hence, the sentence 'unwritten contracts do not bind' is not contradictory (unless of course 'contract' just means *binding agreement*): it does not mean *unwritten agreements that need not be written in order to bind do not bind*, but, rather (something like) *unwritten agreements do not bind*. It is simply false. Presumably Burge would claim that an analogous point applies in the 'arthritis' case: the fact that arthritis cannot occur outside of joints is not constitutive of the concept ARTHRITIS; rather, it is simply not constitutive of ARTHRITIS that arthritis is a disease of the joints only. Applicability of 'arthritis' to ailments occurring outside of joints is undetermined by the term's content. So Bert's utterance of 'I have arthritis in my thigh' is not contradictory or incoherent; it is just mistaken.

But if this is the case, then it's hard to see how "any dictionary could have told [Bert]" that he could not have arthritis in his thigh – how, that is, mastery of the term 'arthritis' would have enabled him to avoid his mistake. If Bert's error is not an ordinary empirical (theoretical) one – because such an error would not drive intuitions in the direction of anti-individualism about the individuation of thought content – then it must be a conceptual one.¹⁶ And, given what Burge says about dictionaries, it's hard to see how Bert's utterance could fail to be conceptually incoherent.

Moreover, the distinction between its being constitutive of a concept *C* that *c*'s are not *F* and its not being constitutive of *C* that *c*'s are *F* cannot serve Burge's purposes. It is not enough that it not be constitutive of the concept ARTHRITIS that arthritis can only occur in joints for Bert's utterance to be conceptually anomalous. Compare: the concept LONELY is not a

constituent of the concept BACHELOR; yet the sentence ‘bachelors are lonely’ is conceptually consonant. Generally, if a concept G (or not- G) is not constitutive of a concept F , it does not follow that the sentence ‘ F is G ’ is in any way conceptually problematic. So to counter Fodor with the claim that POSSIBLY UNWRITTEN is not constitutive of CONTRACT takes the teeth out of the example. If an utterance does not suggest lack of mastery of a concept, then it should come as no surprise that an individual could have a belief of which the concept is a constituent. If there is no conceptual dissonance, there is no reason to think that the match between socially determined linguistic content and mental content is due to anything other than the individual having chosen the correct words.

Now, Burge explicitly disavows (for Quinean reasons) any commitment to a view on which one concept may be a constituent, or be constitutive, of another.¹⁷ But what then does it mean to say that it is *definitional* that F s are not G (that “any dictionary could have told you” that F s are not G)? If there’s nothing like a constitutive connection between a concept F and a concept *not- G* , then how is an utterance of the sentence ‘ F s are G s’ to indicate anything other than a factual error? Perhaps something like a meaning postulate, of the form ‘ $(x)(Fx \rightarrow \neg Gx)$ ’, could provide such a (non-constitutive?) connection. But then any expression of the form ‘ a is F and G ’ (e.g., ‘my ailment is arthritis and is in my thigh’), while not explicitly contradictory, would in combination with the meaning postulate entail the explicitly contradictory expression ‘ a is F and G and not G ’. And anyone who had completely mastered the term ‘ F ’ would know this. So the problem Fodor raises remains: though Bert might not know that his belief involves this conceptual problem, we who ascribe it to him do.

There would seem to be a dilemma here for Burge.¹⁸ If there is no *conceptual* as opposed

to factual problem evinced by Bert's utterance, then the thought experiment does not show anything about the constitution of mental content. On the other hand, if Bert's utterance does involve a conceptual error of the kind Burge relies on, then homophonic ascription violates the principle of charity, since *by the ascriber's lights* the content clause of the ascription is incoherent.¹⁹

Burge does accept that charity precludes ascription of *explicit* contradictions (1982b: 291-92), but he denies that the utterances in his examples are explicitly contradictory. If by 'explicit' Burge means *formal*, then he is surely right. Even assuming that CANNOT OCCUR IN MUSCLE is a constituent of ARTHRITIS, 'arthritis can occur in muscle' – unlike 'a disease that cannot occur in muscle can occur in muscle' – is not a formal contradiction. But neither is 'a disease that can only occur in joints can occur in muscle'. And if it is obvious to the fully apprized user of this last sentence that it is contradictory, then it ought to be obvious to the fully apprized user of 'arthritis can occur in muscle' – in particular, to the ascriber of the belief that he has arthritis in his thigh to Bert – that it is also contradictory. And that ought to be sufficient for the principle of charity Fodor reminds us of to spring the *ceteris paribus* clause of the principle of literal interpretation. All things are not equal in Bert's case because literal interpretation of homophonic ascription violates the presumption of rationality.

Still, Burge could concede to Fodor that the cases presented in "Individualism and the Mental" do not survive his criticism and continue to maintain that social/linguistic anti-individualism is true. For he could dismiss such cases as involving linguistic incompetence (anyone who was really competent with 'arthritis' would not believe he had it in his thigh), and argue on the basis of completely ordinary, unproblematic utterances (as he does in describing

Bert before introducing the problematic belief) that what someone thinks is determined by what they (we) say, given the Principle and the intuitive naturalness of our ordinary ascriptions. He could further claim that part of what it is to be a member of a linguistic community is to be fully competent in the language of that community.^[20]

But we have already seen that non-dissonant examples do not make Burge's case, since there is an obvious alternative explanation of the match between thought content and utterance content. Moreover, it would not help in the case of Al, since it is assumed that he is fully apprized of the meaning of, and hence fully competent in the communal use of, the word 'sofa'. So it can be argued that the principle of charity prevents attribution of the belief that sofas are not furniture to Al since, *by the ascriber's lights*, that belief is conceptually incoherent. Indeed, one could argue that the case as conceived is incoherent even independent of principles governing third-person attributions of thoughts to Al. For, given that Al knows that sofas are by definition pieces of furniture, how could he consistently believe that they are not pieces of furniture?²¹ And how could we, who do not share Al's subversiveness, attribute such a belief to him without impugning his rationality?

Nonetheless, I think the charity objection does not really get to the bottom of the kind of case Burge wants to make for anti-individualism. Given the role the examples are supposed to play in the thought experiments, it really is not necessary that they involve the sort of *prima facie* conceptual *incoherence* Burge's actual examples exhibit. The examples are supposed to illustrate the domineering role of the Principle by presenting it with a serious obstacle to overcome. They are supposed to show that even when an individual's errors raise the specter of reinterpretation, homophonic attribution is completely natural and compelling, and therefore the

Principle still applies. I have argued that the 1979 and 1986 cases run afoul of the principle of charity. But that might only show that Burge's examples were not well chosen. If Burge can accept (because he can explain away) cases of conceptual dissonance in which reinterpretation is forced because there remain so many cases in which it is not, then giving up Burge's examples would not spoil the Burgean argument for anti-individualism if there are others that can play the same role.

The problem with Burge's examples is their conceptual incoherence. But since it is conceptual *dissonance* that is doing the work, and since incoherence and dissonance are not the same thing, there may be examples that make Burge's point without violating the principle of charity. In the next section I consider such dissonant-but-coherent utterances, and argue that, in every sort of case I could think of, though homophonic ascription seems natural and inevitable, there is a clear intuition that literal interpretation is *not* preferred. I think this shows that the naturalness of homophonic ascription and the appropriateness of literal interpretation are subject to different standards, and, hence, that an inference from naturalness to literal truth is invalid.

III. SECOND OBJECTION: MALAPROPS, MISNOMERS AND OTHER ASSORTED ERRORS.

Consider Marge. After surgery, Marge goes on vacation to recuperate. At the airport, she repeatedly sets off the metal detector. Later, Marge says "I began to wonder whether the doctor left a scaffold in me."²² If we say that Marge *said that* she wondered whether the doctor left a scaffold in her, it seems clear that what we are saying is literally true. That *is* what she said – in the sense that this is what the words she uttered mean in the language she speaks. It seems forced to say that in her mouth the word 'scaffold' means *scalpel*, and, hence, that what she said ("the doctor left a scaffold in me") might plausibly be true. Moreover, it seems, intuitively, entirely

natural to *say that* Marge wondered whether the doctor left a scaffold in her. Marge is confused. She thinks ‘scaffold’ is a term for a piece of surgical equipment; but she is mistaken. We describe her confusion using her own words, just as we do in the case of Bert. And after she has been corrected Marge might describe her misstatement using the words she uttered, just as Burge did in the ‘fortnight’ case.

Marge’s utterance is sufficiently dissonant to raise suspicions about her competence with the term ‘scaffold’, and, hence, to present a *prima facie* obstacle to literal interpretation. But it does not run afoul of the principle forbidding casual attribution of contradictory or conceptually incoherent thoughts. It is not at all plausible that Marge’s surgeon would have left a scaffold inside her; but it is not conceptually or logically impossible. (The scaffold would just have to be sufficiently small!) So perhaps this is just the kind of example that Burge needs. It does not involve a contradiction, but neither is it an *ordinary* empirical error. It is a rather *extraordinary* one: scaffolds are not the *kinds of things* surgeons use. It’s *bizarre* to think that a surgeon could leave one in a patient. Anyone who seriously says so thereby provides *prima facie* evidence of not knowing what a scaffold is – of not being competent with the concept SCAFFOLD. However, since the Intuition applies in Marge’s case, the Principle ought to apply as well, and Marge ought to have the concept SCAFFOLD, and think that the doctor left a scaffold in her, in spite of her confusion. The domineering role of the Principle would thus be highlighted, and the social anti-individualist thesis confirmed.

Another sort of example escapes the charity problem, but, arguably, does not involve *conceptual* dissonance at all. Suppose Marge says “Pat Boone was an early American frontiersman.”²³ We would not reinterpret her as having *said that* Daniel Boone was an early

American frontiersman. She may have meant to say something about Daniel Boone, but she did not, because she didn't use his name. Moreover, though the oddness of Marge's utterance might lead us to suspect that she's confused about who Pat Boone is, it is nonetheless intuitively very natural to *say that* she believes that Pat Boone was an early American frontiersman. She made a jarring error; but it's natural to describe it using the very sentence she uttered. And again, if the Principle applies, we can conclude that Marge believes that Pat Boone was an early American frontiersman. Even though she is confused about who Pat Boone is, she still "has" (whatever that means in this context) the name, and may be correctly attributed thoughts using it. What she thinks is determined by how her linguistic compatriots use the name 'Pat Boone'.²⁴

Constructing a Burgean thought experiment using examples like these not only avoids the charity objection; it also comports more comfortably with Burge's Quinean repudiation of a sharp distinction between conceptual truth and empirical truth.²⁵ It might seem that only *constitutive* conceptual errors – i.e., that only errors of (roughly) the form 'a is F', where NOT-F is a constituent of the concept expressed by 'a' – can support Burgean conclusions about mental content. But the examples of this section show that this is not the case. Clearly, ordinary empirical errors don't serve Burgean purposes very well (if at all). If someone says "Aluminum does not conduct electricity" or "Water freezes at 33°F," there's no special reason to doubt that he has the relevant concepts, and no hesitation to overcome in attributing a belief to him with the content of the sentence he utters. It would be *prima facie* pretty implausible to maintain that these errors were due to a misunderstanding of one or more of the terms comprising these sentences. Yet one need not insist that the examples have to involve incoherence or contradiction, or that the concepts involved have analyses, or (perhaps) even that the examples

have to involve confusion about a *concept* at all. A Burgean error just has to be *odd* enough raise suspicion about an individual's competence with a particular term, so that the Intuition can sweep it away and the Principle can show that the individual thinks what he says in spite of his confusion. The Marge examples seem to do this, without running into trouble with the principle of charity.

But there are strong counterintuitions in both sorts of cases. In the malaprop case, if we reflect, and ask ourselves if Marge *really* was thinking that the doctor left a *scaffold* in her (that diamonds have *faucets*, that shoe stores sell *hibachis*, that she took her health for *granite*), I think the answer has to be "No." This really is not what she thought. (She was not thinking what *we* think when we think that Marge's doctor left a scaffold in her.) She simply misexpressed herself. She had a false belief about which word to use to say what she was thinking, and that belief led to her anomalous utterance. In spite of the fact that the Intuition holds, the *ceteris paribus* clause of the Principle is sprung: all things are not equal in Marge's case, because she is not fully competent with the words she uses. She doesn't know what some of them mean.

Now, Burge accepts that malaprops are evidence of a lack of the minimal competence required for an individual to be subject to the Principle (1979: 90). Hence, he ought to agree that it's not literally true that Marge believes her doctor left a scaffold in her. The crucial point here, however, is that it is *nonetheless* intuitively completely natural to say that she believes this. It's as natural as describing Bert's and Burge's errors, and Al's subversion, using their own words. In Marge's case it's clear that we don't accept our homophonic ascription as literally true. We know that Marge meant to say 'scalpel', and that this is what she was actually thinking. We know that she didn't really think the doctor left a scaffold in her. We realize, on (brief)

reflection, that we should not take what it is intuitively natural for us to *say* about her to be literally *true*. So the presence of the Intuition provides no assurance that we should apply the Principle in spite of a speaker's errors. But Burge's argument for anti-individualism depends on this: naturalness of homophonic ascription is supposed to overpower any inclination to reinterpret. Cases of this type show that the Intuition concerns only what it is natural to *say*, since, on reflection, we do not take our ascriptions to be literally true. The Intuition does not prevent the *ceteris paribus* clause of the Principle from being sprung.

There is a strong counterintuition in the misnomer case as well. On reflection, what it's natural to *say* Marge thought is not an accurate description of her. If Marge is *that* confused about Pat Boone, then she really doesn't know who she is talking about, and should not be said to *believe that* Pat Boone was an early American frontiersman.²⁶ Though she may *refer to* Pat Boone by using his name, and though her utterance may express a singular proposition one of whose constituents is Pat Boone, I think it's intuitively natural to say that she was not really *thinking of* him when she said what she said. She was really thinking of Daniel Boone, but used the wrong name to express her thought. If we tell her that 'Pat Boone' is the name of a smarmy mid-20th-century American pop vocalist, and that the name she wants is 'Daniel Boone', she will no doubt say that the latter is who she meant. She did not mean to say (though she did say it) – because she was not thinking – that Pat Boone was an early American frontiersman. Hence, once again, literal interpretation of a spontaneous homophonic attribution is, on reflection, precluded. In both sorts of cases, the Intuition holds, but the Principle does not.

Indeed, I would argue that the Intuition applies in virtually all of the cases Burge explicitly exempts from the Principle.²⁷ Suppose little Francine is taken to the zoo, having been

told that, among other fun things, she will get to ride a big escalator. After passing a number of cages containing large animals, Francine asks, excitedly: “Which one is the escalator? When do I get to ride it?” It seems entirely natural to describe Francine as having mistakenly believed that escalators are animals. This is a perfectly intuitive way to describe the mistake she made, even though hers is a case in which “mastery of the language and responsibility to its precepts have not been developed; and mental content attribution based on the meaning of words uttered tends to be precluded” (Burge 1979: 90). (We routinely describe children’s errors in this way: “When she was little, Lotta thought garter snakes were a kind of clothing”; “As a child, Linus thought guerillas were apes”; “When she was a kid, Nancy thought cars could ride on dental bridges”; etc.) But surely Francine was not thinking what *we* would be thinking if we thought that escalators are animals.

Or consider recently-arrived Andreas. Upon seeing a large Times Square billboard featuring the scantily clad torso of a female supermodel, Andreas exclaims: “Wow! Who is that? She has a really beautiful finger!” Again, it’s quite natural to describe Andreas using his own words – to say that he believed that the woman had a beautiful finger – even if we don’t count him as a linguistic compatriot, and we’re sure that what he was thinking was that she had a beautiful *figure*. He wasn’t thinking what *we* would think if we believed that the woman had a beautiful finger. (This way of describing foreigners’ errors is also quite common: “When Silencio first got here, he thought the Battery was a giant power supply”; “Soon after she arrived, Ludmilla wondered how delis could get away with selling heroin sandwiches”; “Young Sook used to think downtown was a subway stop”; etc.)

The case of spoonerisms is not quite so straightforward. Imagine that after you and

Archie have bought your movie tickets Archie says: “You go get us seats; I want to buy some cop porn at the concession stand.” It does seem fairly natural to report him as having wanted to buy some cop porn at the concession stand, or as believing that movie theaters sold cop porn, even though we know that that’s not what he meant, or what he wanted to do. He made a linguistic mistake – he spoonerized ‘popcorn’ – and we report his error homophonically.²⁸

Burge’s own initial acceptance of literal interpretation in cases of radical misunderstanding shows just how powerful he thinks the Intuitions are. It *is* natural to say that the person who says “An orangutan is a fruit drink” believes that an orangutan is a fruit drink. Our spontaneous impulse is to describe him using the very words he uttered. Burge’s hesitancy with respect to the example, and his subsequent rejection of it, however, suggest that he (at least tacitly) appreciates the gap between the Intuitions and the Principle. We naturally describe the errors of malapropists, misnamers, children, foreigners, spoonerizers and the radically confused homophonically, even while we explicitly exempt them from the Principle. We think that Francine hasn’t the slightest idea what ‘escalator’ means; so we don’t really think she thought that escalators are animals. We *say* she did, and the intuition that our description is appropriate is very strong. But a moment’s reflection shows that we don’t mean it literally. The fact that homophonic ascription is natural in so many cases in which literal interpretation is precluded shows that the Intuitions cannot do the work the Burgean thought experiments require them to do. They do not show that the Principle applies in spite of incompetence. Hence, they cannot be relied upon to make the case that socially constituted linguistic meaning determines psychological content.

IV. THIRD OBJECTION: THE INCOMPETENCE PRINCIPLE.

The burden of the last section was that Burge's reliance on the Intuitions to reveal the hegemony of the Principle – the lynchpin of his argument for anti-individualism – is misplaced. I interpreted Burge as arguing that *because* it's so intuitively compelling for us to use the speaker's own words to describe his mistake, our homophonic attributions are literally true. But this inference isn't valid. Malaprops, misnomers, spoonerisms, the mistakes of children and foreigners, and radical misunderstandings show that the Intuitions can apply in cases where literal interpretation is clearly inappropriate. The strength of the Burgean Intuitions does not make a strong case for literal interpretation or, consequently, for anti-individualism.

Withholding of homophonic ascription in the cases of malaprops, misnomers, etc. may be occasioned by their bizarreness, but it is not based on it. It would be outlandish to think that a doctor left a scaffold in you, or that escalators are animals, or that an orangutan is a fruit drink; but outlandishness is not what makes reinterpretation seem mandatory in these cases. There are plenty of bizarre utterances that do not call for reinterpretation (for example, "I just swallowed my arm," "There's a volcano in my glove compartment," and "My bathroom is made of liverwurst").²⁹ What makes reinterpretation mandatory in these cases is, rather, that the speakers' mistakes indicate a lack of linguistic/conceptual competence: they indicate that the speakers don't know the meanings of some of the words they're using. And that, in turn, is an indication that what they said, according to community standards of meaning and use, is (most likely) not what they meant, or what they were thinking.³⁰ That is, our ordinary attitude-ascribing practices are apparently governed by what I'll call the "incompetence principle":

If a speaker lacks competence with a term *t*, then his *t*-involving utterances should not be taken at face value.

This is a commonsense principle. Ordinary speakers accept it (if only tacitly), and applying it does not require any theoretical knowledge or entail any theoretical commitments. If you don't know the meaning of the words you utter, then you don't mean what they mean, and you're not thinking what they mean. Whatever one's views on the bugbears of contemporary empiricism (synonymy, definition, the analytic-synthetic distinction, conceptual analysis, etc.), if part of being competent with a term *t* (and the concept it expresses) is knowing that it does not apply to *F*s, then anyone who says "*t*'s are *F*" is evincing *prima facie* incompetence with *t*. If the utterance is not a slip of the tongue or other momentary anomaly, then the individual ought to be judged generally incompetent with respect to *t*, and his *t*-containing utterances treated accordingly. He should not be taken to have meant or thought what the *t*-involving sentence he uttered means in his language.³¹

The incompetence principle obviously has application to young children, foreigners, malapropists and the radically confused, as Burge acknowledges (and perhaps to misnamers and spoonerizers as well, if their errors indicate linguistic incompetence³²). In Burge's original cases the mistakes are less startling: they involve relatively subtle errors, in the context of other apparently unproblematic utterances involving the specimen term. But they are sufficient to indicate a lack of competence, and so to make Burge's subjects *prima facie* candidates for reinterpretation.³³ Indeed, Burge *must* accept that the incompetence principle applies in his cases, since the thought experiments require a *prima facie* challenge to literal interpretation: the Principle has to have a counterintuition to overcome if its dominance is to be revealed. The counterintuition is that if an individual doesn't understand a term, he shouldn't be attributed mental states containing the concept expressed by the term on the basis of his utterances of it.

(This is just what it means to say that his utterance should not be taken at face value.) The knowledge Burge's subjects lack, though perhaps minimal, is still crucial for full competence with the concepts expressed by the terms they misuse, and so their lack of it makes them *prima facie* candidates for reinterpretation.³⁴ If they were not, the Intuitions would have no role to play, and Burge's thought experiments would be hobbled. The incompetence principle is one of the engines of Burge's intuitive case for anti-individualism.

But we've seen that the Intuitions don't in general play the role Burge assigns them. They don't override incompetence in the cases of malaprops, misnomers, spoonerisms, children's and foreigners' errors, and radical misunderstandings. The fact that it's natural to *say that* Marge believes that diamonds have faucets, or that Francine believes that escalators are animals, or that Silencio believes that the Battery is a giant power supply, cannot be taken to imply that this is what they believe. Hence, the fact that it's natural to characterize Bert's mistake in the very terms he misuses cannot be a reason to take his utterance (or our ascription) at face value. The Intuitions do not countermand the imperative of the incompetence principle to suspend literal interpretation of Burge's subjects.

Indeed, the failure of the Intuitions to guarantee literal interpretation and the operation of the incompetence principle conspire to short-circuit *any* thought experiment along Burgean lines. An utterance dissonant enough to serve the purposes of the thought experiments will *ipso facto* invoke the incompetence principle: any reason to doubt that an individual has the concept expressed by a word he utters will be a reason to doubt his competence with it, and, hence, will be a reason not to attribute to him the concept expressed by the word. Without the Intuitions to save the day, these reasons are conclusive.

It might be objected that if competence with a concept *C* involves knowing that *c*'s are *F* for some range of predicates *F*, then incompetence can come in degrees (depending on how many of the *F*s one has left out). Bert may be distinguished from malaproposists, misnamers, etc. in that the latter have crossed a threshold of competence beyond which the Principle no longer holds. Bert's utterance evinces minimal (or at least non-egregious) incompetence – in comparison to, say, Marge's or Francine's. There's only a small bit of relatively non-central knowledge about the application of 'arthritis' that Bert is lacking (at least he knows it's a disease that attacks the joints). He hasn't crossed the threshold into gross incompetence (e.g., he doesn't think arthritis is a kind of vegetable), so there is no *prima facie* reason not to interpret him literally. The relevant principle governing ordinary mental state-ascription is that *grossly* or *egregiously* incompetent speakers should not be interpreted literally. There is no intuitive barrier to taking the *minimally* or *subtly* incompetent at their words.

However, as we've seen, exempting Bert, et al. from the incompetence principle sabotages the thought experiments. Moreover, to claim that minimal or subtle incompetence doesn't justify reinterpretation is tantamount to begging the question against the individualist. The intuition that Bert is a candidate for reinterpretation is untendentious. Since he doesn't quite know the meaning of the word he utters, it's natural to suppose that he didn't quite think what he said. This is an individualist intuition. What Bert *said* is determined by the communal content of his utterance, but what he *thought* is determined by factors internal to him (how *he* understands 'arthritis'). Burge argues that this internalist intuition is overwhelmed by the naturalness of homophonic ascription. To assert without argument that the intuition is invalid when incompetence is minimal is to reject without argument the central individualist intuition.

It might also be objected that Burge has another reason not to reinterpret Bert, et al. – namely, that belief attributions are usually not retracted or reformulated when it's discovered that the ascriber has misunderstood a constituent term. Our ordinary practice does not endorse object-level reinterpretation in cases of conceptual confusion like Bert's. This is so because "[i]n many cases, particularly those involving partial understanding, finding an interpretation in accord with the method would be entirely nontrivial" (1979: 94), and any such reinterpretation would be *ad hoc*. But in many cases (e.g., Marge's malaprops) reinterpretation is entirely trivial, and non-*ad hoc*. Moreover, those cases in which it is not (e.g., Francine's) cannot *therefore* be exempted from the requirements of the incompetence principle. If an individual's utterance evinces a lack of competence with a particular term, that individual is not exempt from reinterpretation simply because none is readily available. The unavailability of a reinterpretation may be only a practical matter, whereas the application of the principle is, well, a matter of principle.

In any case, the incompetence principle doesn't require *reinterpretation*: it states only that the communal meaning of an individual's utterance should not be taken to be the content of the mental state he sought to express. This is consistent with not knowing what to say about what, if anything, the speaker really thought, or meant, or was trying to say. Not taking our intuitively natural ascriptions at face value doesn't commit us to providing non-face value reinterpretations of them. Whether or not there is a convenient way, or any way at all, to *reinterpret* incompetent speakers is not in the end an apt consideration for Burge. What really matters is whether or not the speakers can be taken at their words – whether or not they were thinking what they said. They may have thought something else; or they may not have had a (complete) thought at all.

Moreover, in cases such as Bert's in which reinterpretation does seem to be called for, I suspect that the apparent difficulty of finding the right one is due to the belief that if there is no single word to replace the misused one, there is no concept available for reinterpreting the speaker. The assumption that one must find or coin a new term that expresses what the speaker thought is common among Burge's commentators. But there's really no reason to think that reinterpretation must involve term-replacement or neologism. Consider Bert prior to his famous statement to his doctor. He has a painful condition in his joints. He goes to his doctor, who tells him he has arthritis. Let's suppose that Bert had never heard of arthritis, and that it's on this occasion that he learns the word for his condition. (This assumption isn't crucial; it merely makes it easier to state the relevant facts about Bert.) He thus comes to believe that the disease he has in his joints is called 'arthritis'. Some months later, Bert reasons as follows. "The disease in my joints has spread to my thigh. The disease I have in my joints is called 'arthritis'. So now I have arthritis in my thigh." He goes back to the doctor and says "I think I have arthritis in my thigh." Bert's utterance of 'I have arthritis in my thigh' indicates a lack of competence with the term 'arthritis'. Hence, he should not be interpreted as having believed what his utterance literally means. What, then, should we on reflection say that Bert was thinking, if not that he had arthritis in his thigh? If Bert's partial competence with the term 'arthritis' consists in his knowing that arthritis is a disease that occurs in the joints (i.e., if *part of* what it is to be competent with 'arthritis' is to know that it applies to a disease that occurs in joints, and this is *the whole of* Bert's knowledge of its application), but partial competence with a term doesn't ensure possession of the concept it expresses, then it seems reasonable to say that when Bert utters the word 'arthritis' he is expressing the conception that constitutes his partial

understanding. The obvious candidate for reinterpreting Bert (given what we know about him) would then be that he thought that the disease in his joints was occurring in his thigh.³⁵ This attribution violates neither the principle of charity nor the incompetence principle, and it seems to capture exactly what Bert was thinking – what *he* thought he was saying when he said “I have arthritis in my thigh.”

Further, one need not suppose either that Bert thought the term ‘arthritis’ expresses the concept JOINT DISEASE in his language, or that it expresses that concept in his idiolect. All that’s required is that he thought the term ‘arthritis’ applies to the disease he had in his joints, and that that disease (whatever it is) was now attacking his thigh. One need not suppose that he has any concept at all corresponding to the term ‘arthritis’ (if he had had some other disease involving his joints in addition to arthritis – say tendinitis – he might have expressed the very same thought in uttering the sentence ‘I have tendinitis in my thigh’).

Another point Burge stresses is that Bert’s problematic utterance does not prompt us to reinterpret all of his *prima facie* unproblematic ones: we still say without hesitation that he believed he has arthritis in his joints, that arthritis is painful, that arthritis is not as serious as cancer, etc. Reflection on his mistake is not supposed to change any of this. But if the incompetence principle holds, then once an individual has evinced a lack of mastery of a term, *none* of his utterances of it should be taken at face value. If Bert doesn’t quite know what ‘arthritis’ means, then it doesn’t matter whether or not a particular utterance of it exemplifies his misunderstanding. If he doesn’t have the concept ARTHRITIS, then he didn’t believe that he has arthritis in his joints, that arthritis is painful, etc. And if we don’t go back over all our previous ascriptions of arthritis thoughts to Bert, it can only be for practical, not theoretical reasons.

Perhaps in Bert's case there's no point to it.

But it's easy to imagine cases in which there would be a point to it. It might be important that what someone said is not what he was thinking. There may be good reason to reconsider previous attributions made on the basis of ostensibly unproblematic utterances, once the speaker's error is known. Suppose, prior to his visit to his doctor, Bert goes to his pharmacist and says: "I've got arthritis. Can you recommend a non-prescription treatment for it?" The pharmacist recommends an over-the-counter medication, which she tells him to apply to the afflicted joints every morning and evening. Bert says: "Oh no, it's not for my joints – they're fine. It's for the arthritis in my *thigh*." The pharmacist immediately realizes that Bert is confused about what's wrong with him. She's no longer sure he has arthritis. She came to believe that he had arthritis on the basis of her belief that *he* believed he had arthritis; and she came to believe that he believed he had arthritis on the basis of his utterance of 'I've got arthritis'. But now that she knows that Bert doesn't know what arthritis is, the (perhaps unspoken) homophonic ascription she made on the basis of his utterance is suspect. Bert didn't believe that he had arthritis. She may have no ideas about what he was really thinking, but she has evidence that it was not what he said. And since what she thought he believed determined which medication she recommended, and she has an interest in giving him the proper treatment, she has a reason to revisit and re-evaluate her previous attributions.

It's also easy to imagine cases in which reconsideration of past, ostensibly unproblematic utterances is called for where it is obvious how the speaker should be reinterpreted. Suppose Bert is on a spy mission to obtain information about a wealthy foreign sponsor of terrorists, whom his government wants assassinated. Specifically, Bert is to determine where on his large

estate the man will be at various times the following day. Bert reports to his superior, Vivian: “The target is always in the aviary from 3:00 to 4:00 in the afternoon.” Based on what Bert says, Vivian comes to believe that the target will be in the aviary from 3:00 to 4:00 the next afternoon (via her belief that this is what Bert believes). She directs her snipers to be in strategic places around the aviary at the appropriate time. The next morning, Bert says to Vivian: “Remind them to watch out for the bees when they go in to confirm the kill.” “What bees?” Vivian asks. “The bees he keeps in the aviary,” replies Bert. “You fool!” cries Vivian, realizing Bert’s mistake. “An aviary is a place where *birds* are kept, not bees! I’ve sent the snipers to the wrong place!” Vivian has very good reason (viz., the success of the operation) to revisit all of Bert’s previous statements about the location of their target, and to re-evaluate her homophonic attributions. Bert was not thinking that the man would be in the *aviary*, but that he would be in the *apiary*.

Finally, it might be objected that the incompetence principle doesn’t apply in cases like Al’s, in which an individual is fully apprized of the community standards for the application of a term, and is expressing doubts about their correctness. Al knows what ‘sofa’ means in English. When he says “Sofas are not furniture,” he is voicing his suspicion that his compatriots are wrong about what sofas are. Hence, since Al’s utterance does not reflect a lack of mastery of ‘sofa’, the incompetence principle doesn’t apply to him, and so does not prevent literal interpretation.

However, though by hypothesis Al is competent with the term ‘sofa’, his utterance is sufficiently dissonant to moot the issue of reinterpretation for any third party. All things are not equal in Al’s case for his audience, since what Al says will count as *prima facie* evidence that he is confused about the meaning of ‘sofa’ (in addition to being, by any non-skeptical third-party

ascriber's lights, incoherent). Hence, the incompetence principle gives us a reason not to interpret him as believing that sofas are not furniture.

But what of Al's *self*-ascriptions? Burge applies the principle of literal interpretation to third-person ascriptions based on an individual's utterances. This is an effective way to show how communally determined content gains entry, so to speak, into the mental life of an individual: the content of *our ascription* is the content of the speaker's mental state. (Our act of ascribing overrides any individualistic factors that might suggest otherwise.) But it's not the only way. The Principle as Burge states it is not limited to third-person ascriptions ("there is a methodological bias in favor of taking natural discourse literally, other things being equal. ... Literal interpretation is *ceteris paribus* preferred." (88)) So, Burge might as well have focused on self-ascriptions of his subjects.

A first-person Burgean thought experiment would show to an individual who conducts it that the contents of his mental states are determined by the public meanings of the words he uses to self-ascribe them. It would do this by showing him that his self-ascription is intuitively natural and correct, in spite of its dissonance. Al says to himself: "I don't believe that sofas are furniture. They're religious artifacts." This seems to Al a perfectly natural way to formulate his doubt about communal usage, in spite of the fact that he knows that 'sofa' means ... *furniture* ... in his linguistic community. Reflection on this fact could lead Al to see that the content of his sofa-thought is determined by the communally determined content of the sentence he uses to ascribe it to himself. His personal beliefs about the nature of sofas doesn't change the content of his 'sofa' utterances. Likewise Burge, after he realizes his mistake with 'fortnight'. He says (writes) of himself that he used to think that a fortnight is a period of ten days. This seems to

him the intuitively natural way to describe his confusion, in spite of the fact that he now knows that ‘fortnight’ means *period of 14 days*. And in both cases the ascribers are competent with the respective terms, and they know that they are.

Still, the sentences Al and Burge utter to themselves – ‘sofas are not furniture’; ‘a fortnight is a period of ten days’ – are, in the language they speak, conceptually incoherent. As fully apprized speakers they ought to know this, and the principle of charity ought to prevent them from making these ascriptions to themselves.³⁶ They ought to realize that there’s good reason to think that the natural way to describe themselves is not literally true.³⁷ The Intuitions can be as misleading in first-person cases as they are in third-person cases. Moreover, in Burge’s case, *present* competence with ‘fortnight’ doesn’t bring exemption from the incompetence principle with respect to his earlier self. His recognition of his former lack of competence with ‘fortnight’ is sufficient reason for him not to take his homophonic self-ascription at face value.

The fortnight case shows that the incompetence principle applies when one is describing one’s own former incompetence. What is required for a first-person Burgean thought experiment, then, is a competent present-tense self-ascription that doesn’t violate the principle of charity, but which is dissonant enough to provide a *prima facie* challenge to the Principle. An example would be Al saying to himself “I want to buy a new pair of loafers for the shower.” There’s nothing incoherent about wearing loafers in the shower (perhaps Al has an unusual fetish), though Al’s utterance is odd enough to create in a third party some resistance to interpreting him literally. (“Does he really want *loafers*? Why would anyone wear loafers in the shower? Did he mean *loofahs*?”)

The problem with constructing a first-person thought experiment along these lines,

however, is that a speaker's competence, and his awareness of it, deprive the dissonance of his utterance of its essential function for him. Conceptual dissonance is supposed to provide a *prima facie* barrier to literal interpretation, the overcoming of which reveals the dominance of the Principle. But if a self-ascription is dissonant because it's odd (not because it's incoherent), but one knows that one means exactly what the ascription means in one's language, then one has no reason not to take it at face value. The dissonance of the utterance might prompt *others* to doubt the applicability of the Principle; but in the first-person case there is nothing for the Principle to overcome. The result is that such an instance of competent self-ascription will be classed with the ordinary cases described in section I, above. There is no more reason for a competent self-ascriber to think that the content of his thought matches the content of his words because the latter determines the former than there is for him to think that it is because the former determines the latter. AI might recognize that what he says to himself would sound odd to others; but its oddness gives him no reason to doubt that what he said is what he thought.³⁸

V. MENTATION, MENTAL USE, MENTAL MENTION.

We have been led to consider first-person versions of Burge's thought experiment. In an AI-type case, an individual utters to himself a sentence containing a term whose communal meaning he disavows, but nonetheless correctly ascribes to himself thereby a thought whose content matches the communally determined content of the sentence. (AI denies that 'sofa'

expresses the communal concept SOFA, but correctly ascribes that concept to himself with the sentence ‘I believe that sofas are not furniture’.) In a Bert-type case, an individual utters to himself a sentence containing a term he doesn’t completely understand, but nonetheless correctly ascribes to himself thereby a thought whose content matches the communally determined content of the sentence. (Bert has not mastered the communal concept ARTHRITIS, but correctly ascribes that concept to himself with the sentence ‘I believe I have arthritis in my thigh’.) Here, as in the third-person versions, the content of the thought ascribed is the same as the socially determined content of the ascription, in spite of the individuals’ deviance from communal norms.

All of the cases discussed so far have involved both ascription and public utterance; but, in fact, neither is required for a Burgean thought experiment. Burge argues that the content of the words a speaker utters determines the content of his thoughts. But he doesn’t claim that this is so because acts of ascription confer content upon the ascribees’ thoughts. This is just one way to *show that* linguistic content determines mental content: the fact that homophonic ascription is natural and appropriate confirms that the content of the sentence uttered is the content of the thought expressed; that this is so despite the ascriber’s departure from community standards shows that social factors override individual ones in the determination of mental content. It is membership in a linguistic community that establishes the general determinative relationship. An individual’s use of a particular sentence of his language determines the content of the thought he expresses by that use: what he says determines (*ceteris paribus*) what he thinks.

But if it is the use of a sentence of one’s language that confers content upon the thought expressed, then it shouldn’t matter whether the utterance is public or private. If Bert says to

himself (silently or aloud) or to an audience, “I have arthritis in my thigh,” the thought he thereby expresses is that he has arthritis in his thigh, because that is what the sentence means in the language he speaks. As above, we can easily imagine Bert engaging in an interior monologue using the very words he later uses publically to express his belief. If the communal content of what he says aloud determines the content of the thought he thereby expresses, and what he says aloud is the very thing he says silently to himself, then the communal content of what he said silently to himself ought to have determined the content of the thoughts he thereby expressed to himself. It does not seem to be any part of Burge’s argument that a speaker’s utterance has to be actively interpreted by someone else in order for it to have content. Merely speaking, silently or aloud, to yourself or to others, determines the contents of your thoughts.

Indeed, the ‘fortnight’ case of “Belief and Synonymy” seems to depend upon this. For we may suppose that Burge thought through the example and reached his anti-individualist conclusion entirely in his head (or, if not, that he could have). Reflection on his inner utterances led Burge to conclude that the content of his fortnight-thoughts is determined by factors external to him. His claim that the communal meaning of ‘fortnight’ determined the content of his thought in spite of his misunderstanding was based on reflection on an ascription of a belief to his former self (“I used to think that a fortnight is a period of 10 days”), but the content of the ascription itself was not determined by an ascriber (since it is not ascribed). It is determined by the communal meaning of the words that comprise it. Hence, Burge’s inner utterance of the sentence ‘a fortnight is a period of 10 days’ confers upon the thought he thereby expresses to himself the communal content of the sentence.

Moreover, somewhat ironically, we can imagine a Burgean thought experiment conducted entirely internally by anyone who is a minimally competent member of a linguistic community.

Now, genuine first-person versions of Burge's thought experiments (i.e., those that don't collapse into ordinary cases) don't escape the problems of their third-person counterparts.

Interestingly, however, this is not because the Intuitions don't apply in first-person-internal cases. Even if one is fully apprized of the communal meanings of the terms one is inwardly uttering, and even if one recognizes, on reflection, that one's inner utterance is, by community standards, conceptually incoherent, or not in fact what one really meant, it can still seem intuitively perfectly natural and correct to describe one's former errors or present thoughts in the very terms one used or uses idiosyncratically. The Burgean Intuitions, though misleading, are just as powerful in the first-person cases as in the third-person cases. It is an interesting question how this can be, given . I think answering it reveals the ultimate source of the power of the Intuitions.

In note 35, I raised the question how it is that fully competent self-ascribers (Al, Bert and Burge after they have realized their errors) do not recognize that their self-attributions are conceptually incoherent, and that, hence, what they say to themselves is not what they thought. This is essentially the same question as why the Burgean Intuitions are so powerful in the first-person cases.

So we are now led to consider the relation between an individual's inner utterances and the contents of the thoughts he thereby expresses to himself as the most fundamental

Indeed, Bert and Burge are in a position to *remember* that they were not thinking something incoherent. They were not thinking (respectively) that a period of 14 days is a period of 10 days and that a disease that can only occur in joints is occurring in muscle. So what *were* they thinking? And what are they thinking at the moment of their subsequent, fully-informed self-ascriptions?

And Al is in a position to know that he is not thinking something incoherent (and hence, that his attribution is not literally true).

How, if Burge knows that ‘fortnight’ means *period of 14 days*, could he not realize that in attributing to his former self the belief that a fortnight is a period of 10 days he is attributing to himself belief in a contradiction? The beginning of the answer is that the Intuitions are just as powerful in first-person cases as they are in third-person cases: it can seem as intuitively natural to describe one’s own mistakes using the very words one did not understand as it is to describe the mistakes of others homophonically.

The first-person considerations discussed here apply as well to the ascribers in the third-person cases: they are in a position to know that what they are thinking about the ascribees (that, e.g., Bert thinks he has arthritis in his thigh) is conceptually incoherent.

because this is what it's natural to *say*, this is what i *think*

the basic question here is this. when Bert inwardly utters the sentence "I have arthritis in my thigh," is he thereby thinking that he has arthritis in his thigh because that's the content of the sentence he uttered? there's reason to think that he is not, since he doesn't know what 'arthritis' means

the issue is slightly different in the case of Al.

when Al inwardly utters the sentence "Sofas aren't furniture," is he thereby thinking that sofas aren't furniture because that's the content of the sentence he utters? there's reason to think he's not, since, though he knows what 'sofa' means, and so should know that 'sofas aren't furniture' is incoherent, he doesn't think that what he's thinking is incoherent.

Since Burge's principle of literal interpretation requires an utterance to be interpreted, if there is to be a first-person Burgean case for anti-individualism it must involve either soliloquy or inner speech. The individual would have to both recognize the dissonance of his self-ascription and have the intuition that, nonetheless, it is the natural way to think about what he is thinking. Of the two modes of self-ascription, the more basic is internalized utterance. For, even in cases where an individual is speaking out loud to himself, or out loud to another, his outer utterance may be a reproduction of, or an accompaniment to, an inner one.

And just as there are important questions to be asked and answered about the relation of thought and outer speech, there are equally important questions to be asked and answered about the

relation of thought and inner speech.

If an internalized version of the Burgean thought experiment can be run, this just shows that there is a difference between thought and inner speech, since the experiments require that there be the possibility that what an individual has said is not what he thought. If this possibility obtains with respect to inner speech, then it must be admitted that thought is not the same as, though it may frequently or even inevitably accompany, inner speech.

Bert is not expressing the thought that he has a disease called ‘arthritis’ in his thigh when he says “I have arthritis in my thigh.” But it doesn’t follow that he is therefore expressing the thought that he has arthritis in his thigh. What he might be doing is reproducing out loud an inner utterance of the sentence ‘I have arthritis in my thigh’ – an inner tokening, an inner use, of the phonological type. An inner tokening of ‘I have arthritis in my thigh’ is a different event from an inner tokening of ‘I have a disease called “arthritis” in my thigh’. The former uses the term ‘arthritis’ and the latter mentions it. Moreover, neither event can be identified with thinking that he has arthritis in his thigh. One can inwardly token the words ‘I have arthritis in my thigh’ without thereby thinking what they express in one’s language. Hence, vocalizing an inner utterance is not the same as expressing a thought. So the mental state Bert is expressing, though it might not be *metalinguistic*, might nonetheless be *linguistic*, in the sense that what he says to his doctor is merely a broadcasting of an inner utterance (of a sentence he does not entirely understand).

One can imagine Bert reasoning as follows. “The doctor used the word ‘arthritis’ to describe the disease in my wrists and ankles. So the disease I have in my wrists and ankles is called ‘arthritis’. Therefore, I have arthritis in my wrists and ankles. The disease that I have in my wrists and ankles has spread to my thigh. So the condition I have in my thigh is called ‘arthritis’ as well. So now I guess I have arthritis in my thigh.” Having moved the thought experiment inside Bert, we can ask the same questions we asked about the externalized thought experiment. When Bert reasons from his metalinguistic belief to an object-level belief, the intuition is that the content of the word mentioned in the former is transferred to the latter. But this intuition itself depends on the assumption that thought content is determined by socially individuated meanings, and in this case there is no principle of belief-ascription to do the work of establishing the identity. So what justifies Bert’s inference from “The disease in my wrists and ankles is called ‘arthritis’ to “The disease in my wrists and ankles is arthritis”?

If we assume that thought just *is* inner speech, the connection is guaranteed. In fact, I think that something like this must be what Burge has in mind. For the parallels with external speech are suggestive: just as you can say things by publically uttering words you don’t understand, you can think things by privately uttering words you don’t understand.

We can agree with Burge that Bert’s utterance of “I have arthritis in my joints” is not an expression of the metalinguistic thought that the disease he has in his joints is called ‘arthritis’, but still recognize the importance of the metalinguistic belief in the etiology of his utterance.

Though it is not made explicit in the above story, if Bert's utterance depends on his reasoning, and his reasoning depends on his belief that the disease in his ankles is called 'arthritis', then his utterance depends on, though it doesn't express, the metalinguistic belief. And if one's utterance is based on a false belief about the words one utters, then the *ceteris paribus* clause is sprung. (This consideration would apply in cases where Fodor, et al.'s would not – e.g., “ice is inflammable.”)

In general, one can inwardly utter words without thinking what those words express *if one doesn't know the meanings of the words*. This is true in the case where one doesn't know the meaning of *any* of the words one internally utters and in the case where one doesn't know the meanings of *all* of the words one internally utters. So merely uttering the words internally, whatever their social meaning, is not in itself thinking what those words mean. Therefore, even if, were one to utter those words, one would be literally interpreted, it doesn't follow that the literal ascription would be true, if one doesn't know the meaning of one of the words one utters.

if self-ascription is to fall under the principle that *cp* literal interpretations of *an individual's words* are preferred, then self-ascriptions ought to involve internal utterances. [note, this is not a case simply of thinking an object-level thought; it's a self-*ascription*. [But can't one self-ascribe *without* using words? Yes. (Burge would describe it as a self-*attribution*.) But then the content of one's self-attribution is the content of the words one *would* utter in expressing it. Burge can avoid inner utterance using the counterfactual principle. However, if it's possible to choose the wrong word to express what one is thinking, then the counterfactual principle can't be taken to apply universally – either in public or in private.

[Al] the intuition that thought is (or is determined by) inner speech. Just as one can say things using words one doesn't understand, one can think things using words one doesn't understand. It seems clear, however, that, in general, thought is no more inner speech than it is outer speech. For, though it may be true that most of us think *with* words, since it's possible to utter words inwardly without thinking *anything*, it's possible to utter words inwardly without thinking what they mean in the language one speaks. Inner speech is not the same thing as thought. So though Bert *says to himself* "I have arthritis in my thigh," since he doesn't know what 'arthritis' means he no more thinks the thought that sentence expresses than he understands the sentence when he utters it.

"There is nothing irrational or stupid about the linguistic or conceptual errors we attribute to our subjects. The errors are perfectly understandable as results of linguistic misinformation." (100)

I think the only way to make sense of Burge's claim that Al's self-description (and, by parity, Burge's own self-description in Burge 1978) is not in fact incoherent is to suppose that Al's thought that sofas aren't pieces of furniture is *just* an inner utterance of the sentence. For since the sentence itself is not explicitly incoherent, the only way to allow Al to think that sofas aren't pieces of furniture without thinking that pieces of furniture are not pieces of furniture (and to allow Tyler to think he used to think that a fortnight is a period of 10 days without thinking that he used to think that a period of 14 days is a period of 10 days), is to claim that thinking that

sofas are not furniture is just to inwardly token (*use*, not mention!) the sentence ‘sofas are not furniture’.

VI. THE PRAGMATICS OF HOMOPHONIC ASCRIPTION.

VII. CONCLUSION.

NOTES

1. In cases where relations to physical environment have a role in determining content, such relations can also be mediated by social environment (Burge 1982: 102).
2. Though I begin with the arthritis argument of Burge 1979, the issues I raise apply equally to all of Burge's thought experiments, as will become clear in the subsequent discussion.
3. As Burge says (78): "In actual fact, 'arthritis', as used in his community, does not apply to ailments outside joints. Indeed, it fails to do so by a standard, non-technical dictionary definition."
4. In general:
 - a. The meanings of linguistic expressions are determined socially (linguistic anti-individualism).
 - b. The contents of an individual's thoughts are determined by the linguistic meanings of the words he does or would utter to express his thoughts. Hence,
 - c. The contents of an individual's thoughts are determined socially.
5. Burge's claim (83 (my emphasis)) that "the thought experiment does appear to *depend on* the possibility of someone's having a propositional attitude despite an incomplete mastery of some notion in its content" is potentially misleading. Both the anti-individualist conclusion and the claim about incomplete mastery *follow from* the literal truth of premise (1) (in conjunction with the assumption of linguistic anti-individualism). If (1) is literally true, then the content of its that-clause is the content of Bert's belief. But if Bert believes that he has arthritis in his thigh, then he has the concept ARTHRITIS, despite his misconception. The thought experiment depends on the possibility of possession without mastery not as a presupposition, but as an entailment: if possession without mastery is impossible, then Bert can't believe he has arthritis in his thigh, and there must be something wrong with the intuition that homophonic attribution is appropriate.
6. I assume that the thesis of *linguistic* anti-individualism is uncontroversial in this context. Perhaps another way to challenge Burge would be to insist that the meanings of an individual's words are determined by the contents of the intentions with which he utters them. In granting Burge this thesis, the present challenge to his thought experiments is the stronger one.
7. Burge recognizes this: "communal practice is a factor ... in fixing the contents of my attitudes, even in cases where I fully understand the content." (85)

8. That AI may be correctly characterized as challenging the very practices that determine the communal meanings of the words he uses (Burge 1986: 707) does not prevent those practices from determining the content of his challenge. If it did, then the thought experiment would not show that his thought contents are extrinsically determined: they would be determined by what he thought sofas were.

9. Cf. Evans XXXX: xxx.

10. It is not entirely clear how these categories are to be distinguished. A malaprop seems to be more than a slip of the tongue (which is typically corrected once the speaker realizes what word(s) he uttered) but less than a radical misunderstanding of the nature of the term's referent (there is a difference between thinking that 'orangutan' is a word for a fruit drink and thinking that an orangutan is a fruit drink, even if Burge is right). I'll assume here that a malaprop involves a mistaken standing belief about what something is *called* (which word one is supposed to use to refer to it). Whether or not this entails a (radical) misunderstanding of the word's referent, and, if so, whether or not any anti-individualist conclusion can be inferred therefrom, are separate issues. I intend the characterization of malaprops here to be neutral on these questions.

11. Burge has in fact (in conversation) disavowed the intuition about the orangutan case.

12. I am of course not suggesting that Burge thinks otherwise.

13. See also Bilgrami 1993 and Wikforss 2001. Fodor's paper has languished in obscurity for over 20 years. (It is cited by neither Bilgrami nor Wikforss; and Fodor confessed, in conversation, that even he had forgotten it.) Of course Fodor went on to accept Burge's argument (if only out of curiosity; see Fodor 1987: xx) and to develop the by now standard distinction between "narrow" and "broad" (or "wide") content. It seems to me, however (as it does to Bilgrami), that this objection is decisive with respect to the thought experiments in "Individualism and the Mental" and "Intellectual Norms and Foundations of Mind."

14. Whether one can believe a contradiction or not is controversial. See e.g. Marcus 1983, 1990 for arguments that one cannot. In any case, it must be possible to *think* a contradictory proposition. (Otherwise, how could it be *disbelieved*?)

15. Perhaps it is due to inattention, or laziness, or the exigencies of social/linguistic interaction, or the fact that it is not obvious how, given his error, the speaker's belief should be characterized. For whatever reason, we do (as I will argue below) routinely use object-level sentences to report what we take to be mistaken meta-level beliefs. This is as much a part of our ordinary practice as assuming (all things equal) rationality, linguistic competence and truthfulness.

16. An ordinary empirical error would not provide the *prima facie* reason for thinking that Bert does not have the concept ARTHRITIS – the conceptual dissonance – the overcoming of which by the naturalness of homophonic ascription shows the "domineering" role of the Principle.

17. He also says (1979: 88) that his “arguments remain plausible under any of the relevant philosophical interpretations of the conceptual-ordinary-empirical distinction.” The present objection, however, shows that this is not the case. On the traditional decompositional approach to concepts and concept-possession, the fact that ‘arthritis in my thigh’ is not an explicit formal contradiction is a superficial and irrelevant one.

18. Cf. Wiksfors 2001: 221-26.

19. Burge’s (1989/1993) distinctions between translational meaning and explicational/lexical meaning, and concepts and conceptual explications/conceptions will not help here, for two reasons. First, his rationale for the distinctions relies on the results of the thought experiments, and, second, if the latter sorts of contents are tailored to avoid attributing inconsistent beliefs, then either there is no match between ascribers’ and ascribees’ content, or there is, but it is not *in spite of* ascribees’ errors. (For a different response to Burge on these distinctions, see Bilgrami 1992: 39-41.)

20. [someone does argue this – Ebbs? Woodfield? Antony?]

21. I think the case of Al (and Burge’s description of himself in Burge 1978), which involves *self*-ascription of content, does show something deep and important about thought – though not what Burge thinks it does. I have more to say about this, as well as why I think Burge maintains that the thought that sofas are not furniture is consistent, below.

22. It’s not a slip. When pressed (“She left a *scaffold* in you?”), Marge does not retract her statement.

Marge, an actual person, is a rich source of malaprops. She has also been known to say “You can judge the quality of a diamond by how many faucets it has,” “My husband’s shoe store sold fifty pairs of hibachis in one week!” and “When I was young I took my health for granite.” So, for Marge at least, this sort of mistake is “legion.”

23. She didn’t really say this. She did say, however, that her favorite ballet was *Swan Song*.

24. This sort of example suggests that what is going on in the other cases might not involve conceptual contents either. I’ll have more to say about this below.

25. Though one might insist that taking non-contradictory but very bizarre utterances at face value does violate the principle of charity, more broadly construed: one should not attribute beliefs that impugn the ascriber’s common sense. (A sensible person would not suspect that a doctor had left a scaffold in him.)

26. This sort of case is different from Evans’s case (Evans 1982: xx) in which an individual successfully thinks about and refers to an individual (Louis) whom he has never met and knows nothing about, simply by picking up his name from the conversation of others. If one simply acquiesces in the practice of those from whom one gets a name, then – at least arguably – one

may think thoughts whose content is determined by that practice. But Marge attaches a number of substantively erroneous beliefs to the name 'Pat Boone'. There is evidence that she is confused about who he is. If competence with 'Pat Boone' requires not believing that he's an early American frontiersman, then Marge hasn't mastered the name.

27. It might not apply to slips of the tongue that are immediately corrected by their utterers.

28. My intuitions about spoonerisms are not as strong or stable as those in the other sorts of cases. Surely it's completely appropriate to report that Archie *said that* he wanted to buy some cop porn. But there seems to be an important difference between utterances that express or implicate mistaken standing metalinguistic beliefs and utterances that, like perhaps spoonerisms and certain slips of the tongue, involve only momentary lapses. The former appear to support the Intuition much more solidly. (At least in my own case. It is more difficult, though not impossible, for me to hear homophonic mentalistic ascription as natural for spoonerisms, especially egregious or bizarre ones. If (as I think is the case – I develop this claim in more detail below) object-level mentalistic ascriptions are often used to report metalinguistic beliefs ('Francine thought that escalators are animals' can be used to report Francine's belief that 'escalator' is a word for a kind of animal) and to report the contents of *utterances* ('Andreas thought the woman had a beautiful finger' can be used to report that the English sentence Andreas uttered means that the woman had a beautiful finger), as well as to report the contents of thoughts, then we could describe the content of Archie's utterance in two ways. We could say either "Archie said 'I want to buy cop porn at the concession stand'" or "Archie wanted to buy cop porn at the concession stand." Since it is unproblematically true that (in the relevant sense) Archie *said that* he wanted to buy cop porn at the concession stand, homophonic ascriptions used in this way could be appropriate and true (though not literally). (We could view them as pragmatically expressing a proposition which is not their literal linguistic content.)

In cases where an utterance expresses or implicates a standing metalinguistic belief, an utterance of an object-level description could be doing double pragmatic duty, reporting both a metalinguistic belief and an utterance content. If, however, spoonerisms don't express or implicate standing metalinguistic beliefs, one of the pragmatic functions of object-level ascription would not be served. It might, therefore, be intuitively less natural to make them, and harder to hear them as appropriate, in those cases.

29. Cf. Davidson 1984. I won't try to figure out here exactly what it takes to bring an individual's conceptual competence into question.

30. It's *possible*, though unlikely, that an individual might not know what a term *t* means in his language but nonetheless, by coincidence, express a thought containing the concept expressed by *t* by uttering a sentence containing *t*. Let's say Mike can never remember the difference between inversion and obversion of a categorical proposition. Though he knows that one of them is replacement of the predicate term with its complement plus reversal of the quality of the proposition and the other is replacement of subject and predicate terms by their complements, he doesn't remember which is which. It seems reasonable to say that he's not competent with the

terms ‘inverse’ and ‘obverse’. He doesn’t know how to apply them correctly. He might, however, intend to say that a proposition p is derived from a proposition q by replacing the subject and predicate terms of q with their complements, and just make the lucky guess that propositions so related are inverses. So when Mike says “ p is the inverse of q ,” what he says is true, what he says is what he thought, and what he says is what he meant to say – though, since he lacks competence with ‘inverse’, it would be a mistake to take his words at face value. His utterance is not a legitimate basis for attributing to him the belief that p is the inverse of q .

31. Burge argues that metalinguistic reinterpretation of Bert, et al. is inappropriate. Bert’s utterance “I have arthritis in my thigh” is an expression of the belief that he has arthritis in his thigh, not the belief that he has a disease called ‘arthritis’ in his thigh. But Bert does have beliefs about the term ‘arthritis’, and they figure prominently in the etiology of his utterance. (Cf. Crane 1991: xx.) As Crane points out, an individual’s choice of words in producing an utterance is determined, in part, by what he thinks they mean. Bert has the true belief that ‘arthritis’ applies to the disease he has in his joints, and the false belief that it applies to the condition in his thigh. If, however, one’s utterance is based on (or caused by) a false belief about the meaning of a word one utters, then the incompetence principle applies to one’s utterances of it. It doesn’t matter that the utterance that reveals the mistaken metalinguistic belief is not itself a direct expression of it. Bert’s mistake is evidence that he has a false belief about what ‘arthritis’ means, and it’s that which invokes the incompetence principle.

32. If they don’t, then reinterpretation of slips and spoonerisms would be covered by a separate principle, e.g.: if a speaker’s utterance involves a malfunction of some speech processing mechanism, then that utterance should not be taken at face value.

33. Recall that Burge’s thesis that partial mastery is consistent with full possession is a consequence, not a presupposition, of his thought experiments.

34. Or, if reinterpretation is not so obvious as replacing ‘scaffold’ with ‘scalpel’ or ‘finger’ with ‘figure’, at least balking at interpreting them literally. (It’s doubtful that there even is any object-level reinterpretation of Francine’s utterance that would capture what she was thinking. In her case it seems the problematic utterance is based on a metalinguistic belief – that ‘escalator’ is a word for a kind of animal – alone. She was not expressing any object-level thought at all.)

35. Or, more awkwardly, that he thought that he had the disease he has in his joints in his thigh.

36. An utterance of ‘I mistakenly thought that aviaries are places where bees are kept’ might (if it is implicit that aviaries are places in which *only* birds are kept) be in the same boat as ‘I have arthritis in my thigh’ and ‘sofas are not furniture’. Marge’s utterance of ‘I used to think a scaffold was a scalpel’ would be even more egregious.

37. Indeed, they are in a position to *remember* that they were not thinking something incoherent. So why don’t they recognize that their self-attributions are conceptually incoherent? And what were they really thinking at the moment of their self-ascriptions? I address these important

questions in the next section.

38. What if Al really did mean to say 'loofahs' to himself, and but said 'loafers' instead? If the mistake is a mere slip, it seems unfair to hold Al (and for him to hold himself) to having thought what he said. If it's a standing error, then he's not a competent speaker.