

## Expression, Thought, and Language\*

### 1. Introduction

It is natural to think of our sincere utterances as expressions of our beliefs. If I sincerely utter “Baseball is popular in Boston,” it’s probably because I believe that *baseball is popular in Boston*. Sincere speakers say what they think, and this suggests the following *expressive constraint* on accounts of the relation between thought and utterance content.

(EC) When a speaker expresses a belief with a sincere utterance, the utterance and the belief have the same content.<sup>1</sup>

The expressive constraint effectively requires that whatever determines either thought or utterance content will determine both. However, the two most obvious accounts of the relation between thought and language compatible with the constraint (giving an independent account of linguistic meaning and explaining thought content in terms of it, and giving an independent account of thought content and understanding linguistic meaning in terms of it) both face serious difficulties. Because of this, the following will suggest an alternative picture of the relation between thought and language that remains compatible with the constraint. Such an account will stress the interdependence of belief content and linguistic meaning, an interdependence that comes from the fact that our language is itself one of the things about which we have many beliefs.

### 2. Should we really endorse the constraint?

Nevertheless, there is no need to make one’s account of the relation between thought and language compatible with the expressive constraint unless it really is something that needs to be accommodated. After all, the plausibility of the expressive constraint might trade on our mistaking an *empirical generalization* for a conceptual truth. That is to say, the supporters of the expressive constraint may endorse (EC) when all our actual practice supports is:

---

\* I’d like to thank Robert Brandom, Joe Camp and John McDowell for comments on earlier drafts of this paper, and the Canada Council for their generous support.

<sup>1</sup> Forms of the expressive constraint are explicitly endorsed in Searle 1983, p. 164, McGinn 1982, p. 217, Bilgrami 1992, p.1. Though, as will be discussed in the following section, an implicit commitment to it is extremely widespread.

(EG) When a speaker expresses a belief with a sincere utterance, the utterance and the belief *typically* have the same content.

While the contents of our words *typically* reflect the contents of our thoughts, this need not be because of the sort of ‘internal connection’ that the expressive constraint postulates. Thought and utterance contents may typically correspond because, by and large, we know what we are thinking, we know what our words mean, and we make an effort to see that the sentences we utter correspond to the contents of our thoughts. Such an explanation requires not an ‘internal’ connection between thought and utterance content, but rather a more ‘accidental’ (though by no means coincidental) one which would break down when, say, we make slips of the tongue or have an imperfect mastery of our language. As Evans notes, “We are all familiar with cases in which, though carelessness or ignorance of the language, the speaker selects words unsuitable to his thoughts.”<sup>2</sup> Dummett expands on this apparently familiar thought as follows:

When an utterance is made, what the speaker *says* depends upon the meanings of his words in the common language; but, if he thereby expresses a belief, the content of that belief depends on his personal understanding of those words, and thus on his idiolect ... In unhappy cases, therefore, his words, understood according to their meanings in the common language, may not be the best expressions of his belief, or may even misrepresent it.<sup>3</sup>

If Dummett and Evans are right about this, the expressive constraint is at odds with our intuitions about a large range of cases. For instance, when Archie Bunker says “We need a few laughs to break up the monogamy,” we attribute to him the belief that a few laughs are needed to break up the *monotony*, so the conventional meaning of what he says does not determine the belief we attribute to him.<sup>4</sup> In much the same way, if someone (hereafter “Bert”) claims “I’ve got arthritis in my thigh,” we can treat him as making a false claim about arthritis (which can’t occur in the thigh), while attributing to him a true belief about, say, *tharthritis* (an ailment including both arthritis and pains in the thigh).<sup>5</sup>

---

<sup>2</sup> Evans 1982, 67-8.

<sup>3</sup> Dummett 1991, 88. See also Dummett 1993, 144.

<sup>4</sup> Indeed, much of the humor of Bunker’s utterances would be lost if we merely attributed to him the beliefs corresponding to the conventional meaning of his sentences. See Davidson 1986 for a discussion of this case.

<sup>5</sup> This last claim is, of course, much more contentious. Bert’s usage (and the denial that we should interpret his thoughts this way) is discussed in considerable detail in Burge 1979, 1989. For a discussion of how related sorts of ‘semantic externalism’ should lead one to reject the expressive constraint, see Lewis 1979, p. 143.

Such cases might suggest that thought and utterance content correspond only when certain empirical preconditions are met. We can see the role of such preconditions by considering Kripke's "disquotation principle" which claims that: "A normal English speaker who is not reticent will be disposed to sincere reflective assent to '*p*' if and only if he believes that *p*."<sup>6</sup> While the disquotation principle initially sounds very much like the expressive constraint, the principle is meant to be about "normal" and "reflective" speakers, and Kripke elaborates these qualifications as follows:

When we suppose that we are dealing with a normal speaker of English, we mean that ... he uses the sentence to mean what a normal speaker of English would mean by it.... The qualification "on reflection" guards against the possibility that a speaker may, through careless inattention to the meaning or other momentary conceptual or linguistic confusion, assert something he does not really mean, or assent to a sentence in linguistic error.<sup>7</sup>

Qualified in this way, the disquotational principle posits a connection between thought and utterance content for neither those who make slips of the tongue or malapropisms, nor those who, like Bert, have an imperfect mastery of the public language. If Kripke's disquotation principle really captured all there was to our pre-theoretical conception of the relation between thought and utterance content, then there would be no reason to see it as requiring the 'internal' connection between thought and language favored by (EC). Rather, the connection between the two would, as suggested by (EG), hold true only in certain specified circumstances.

However, the defender of the expressive constraint can resist such apparent counterexamples to (EC), by questioning the status Dummett's initial claim that "when an utterance is made, what the speaker *says* depends upon the meanings of his words in the common language."<sup>8</sup> The assumption that cases like Bert and Bunker are counterexamples to the expressive constraint presupposes that the connection *between conventional meaning and utterance content* is an 'internal' one. Consequently, the defender of the expressive constraint can argue that it is precisely *this* connection between utterance content and conventional meaning that should be

---

<sup>6</sup> Kripke 1979, 113.

<sup>7</sup> Kripke, 1979, 113.

<sup>8</sup> This thought of Dummett's is put in an especially vivid form in Evans 1982, 67-9.<sup>9</sup> Furthermore, in addition to philosopher's (often implicit) commitment to the expressive constraint, most linguists view the 'conventional' meanings of a sentence as itself determined by how it is *typically* used. As a result, they deny that any *particular* use of a sentence must have the conventional meaning associated with it.

thought of as breaking down on occasion. Dummett may be right to claim that a speaker's words, "understood according to their meanings in the common language, may not be the best expressions of his belief," but this may simply be because in some cases his words should not be "understood according to their meanings in the common language." The content of our utterances may frequently correspond to the conventional meaning of the sentences uttered, but they need not always do so. We can preserve the connection between thought and utterance content by loosening the connection between the contents of utterances and the contents conventionally associated with utterances of their form. The claim that an individual's words have the meanings conventionally associated with them may, then, merely be an *empirical generalization* that breaks down in cases like Bunker and Bert.

The undeniable existence of an occasional gap between the contents of our thoughts and the conventional meaning of what we say thus does not automatically entail that we should reject the expressive constraint. The question is not whether there can be a gap between what we believe and the contents conventionally associated with what we say, but rather where this gap should be located. The defender of (EG) wants to put it between thought and utterance content, while the defender of the expressive constraint wants to put it between utterance content and conventional meaning. However, there are good reasons for putting the gap in the latter of the two positions.

The suggestion that we should loosen the connection between conventional meaning and utterance content is not, after all, merely an *ad hoc* attempt to preserve the expressive constraint in the face of apparent counterexamples. Indeed, such a break between conventional meanings and utterance content follows directly from most philosophical accounts of meaning.<sup>9</sup> For instance, 'classical' accounts of meaning, in which the meaning of a word is determined by the idea associated with it, clearly tie utterance content directly to thought content and only accidentally to conventional meaning.<sup>10</sup> If words get their meaning by being labels for 'ideas' in the mind, then the content of our thoughts and utterances (being combinations of our ideas and words) will be the same. While few people accept the classical account of meaning, the general

---

<sup>10</sup> Consider, for instance, Hobbes, 1651, p. 101, Locke 1689, Book III, Chapter II.

<sup>11</sup> See Grice 1989, Schiffer 1972, Bennett 1976, Appiah 1985, Sperber and Wilson 1986.

strategy of explaining the content of our utterances in terms of the contents of our thoughts can be found in much contemporary work. For instance, ‘Gricean’ accounts of meaning tie the belief contents associated with our communicative intentions to the contents of our utterances.<sup>11</sup> One could also view generative semantics and at least some versions of the ‘language of thought’ hypothesis as committed to a similarly tight connection between thought and utterance content, with only an indirect connection between utterance content and conventional meaning.<sup>12</sup> Finally, a clear and insistent case for a separation between the content of what we say and the meanings conventional conventionally associated with our words is found in the work of those on the ‘Davidsonian’ tradition.<sup>13</sup> The fact that most philosophical accounts of meaning commit themselves to the expressive constraint does not, of course, entail that it must be accepted. Nevertheless, the fact that the constraint is entailed by most attempts to give a systematic account of thought and language (attempts which can be very different otherwise) suggests that the expressive constraint may lie very close to the heart of our presystematic semantic intuitions.

Furthermore, replacing the ‘internal’ connection between thought and language with a more accidental one would not come without certain conceptual costs. This can be seen when we consider some of the consequences of prying utterance content apart from belief content and tying it to conventional meaning. For instance, even though I can’t tell mosquitoes apart from some other flying insects, my thoughts of the form *my word “mosquito” refers to mosquitoes* seem obviously true. If the expressive constraint is given up, however, such ‘obvious truths’ could turn out to be false. Thoughts of the form *my word “mosquito” refers to mosquitoes* need not be true, because what determines the content of my mosquito-thoughts is not the same as what determines what my word “mosquito” refers to.

A related, and more serious, consequence of giving up the expressive constraint would be that one’s self-ascriptions of belief would often turn out to be false. If one has a less than perfect mastery of the conventions associated with the terms in one’s language then, even without

---

<sup>12</sup> See Fodor, 1975.

<sup>13</sup> See especially Davidson 1982, 1986 and 1994. See also Bilgrami 1992 for a development of this aspect of Davidson’s thought.

malapropisms and slips of the tongue, the content of one's *belief* that *P* (hereafter *PI*), which is determined by one's *understanding* of the conventional meanings, need not correspond to the content of one's *utterance* "*P*" (hereafter *P2*), which is determined by the conventional meanings themselves. As a result, if one makes a self-ascription of the form "I believe that *P*," the content of the '*P*' in one's self-ascription is *P2*, but the content of one's belief is *PI*. One may not *believe P2* at all, so the self-ascription may turn out to be false. Some might be willing to bite this bullet for malapropisms and even cases like Bert's use of "arthritis", but it should be noted that the problem would be more widespread than this. For instance, if I know nothing that distinguishes Beeches from Elms, then it may seem as my 'elm' sentences will not reflect my 'elm' beliefs, which cannot help themselves to the more determinate conventional meanings that my utterance content helps itself too.<sup>14</sup> Our self-ascriptions would not be true whenever our own mastery of the public language is less than complete. However, the claim that sincere assertions of mine such as "I believe that there are more than 20 Elms in the United States" are *false* seems very hard to swallow.<sup>15</sup> Since we are often aware that our linguistic mastery is not complete, one would have to posit a serious lack of self-knowledge to explain our mistaken belief that such ascriptions are true.

Perhaps the most serious conceptual problem with giving up the expressive constraint is that it is intimately connected to the intuitions behind Moore's paradox.<sup>16</sup> If one gives up the expressive constraint, one can no longer make sense of what is supposed to be so 'paradoxical' about 'Moore sentences' of the type "*P* but I don't believe it." For instance, if thought and utterance contents differ, then one might frequently be entitled to make assertions such as "Positrons have mass, but I don't believe it." After all, while one might believe that the public truth/assertion conditions (whatever they may be) of "Positrons have mass" are satisfied, one might also believe that (because of one's half-baked knowledge of physics) the content of one's

---

<sup>14</sup> For a discussion of this case, see Putnam 1975. Dummett is apparently willing to bite the bullet on this issue (see Dummett 1993, p. 144).

<sup>15</sup> Of course one would also have to claim that 3<sup>rd</sup> personal ascriptions such as "Putnam believes that there are more than 20 elm trees in the United States" would also be false whenever the subject of the ascription failed to have a full mastery of the terms involved.

<sup>16</sup> See Moore 1993.

own belief “positrons have mass” will not correspond to the content of the assertion made in the public language. While some have denied that our self-ascriptions need all be true, the infelicity of the Moore sentences has never been questioned, and this is an important ‘expression’ of our commitment to the expressive constraint.

These may not be conclusive reasons for explaining the occasional differences between conventional meaning and belief content in terms of a gap between utterance content and conventional meaning rather than one between belief and utterance content. Nevertheless, they do give us some *prima facie* reasons for thinking that we should keep it unless it leads us to some fairly serious conceptual difficulties elsewhere. Of course, as will be discussed below, such difficulties may come to light when we try to describe the relation between thought and utterance content in a way that accommodates the constraint.

### **3. Accommodating the constraint**

Even if we loosen the connection between utterance content and conventional meaning, one might think it easier to account for thought and utterance content if one did not have to accommodate the expressive constraint. After all, there are a number of strong intuitions tied to the content of what we believe (relating to the explanation of behavior, self-knowledge, supervenience, etc.) that we do not feel (or at least do not feel as strongly) with respect to the content of what we say. In much the same way, there are a number of strong intuitions about what we mean by our words (meaning is conventional, public, shared, etc.) that are not felt as strongly about the contents of our beliefs. These sets of intuitions do not always seem compatible with each other. Consequently, if we were to look for independent accounts of belief and utterance content, the two resulting tasks might be considerably simpler than the task of finding a single account that worked for both.<sup>17</sup> Nevertheless, once utterance content and

---

<sup>17</sup> Dividing the subject matter of one’s theory of content this way seems more ‘natural’ than dividing it by giving separate accounts of ‘wide’ and ‘narrow’ content. (See, for instance, Field 1978, Putnam 1981, McGinn 1982, Loar 1985, Block 1986.) The thought/linguistic content split corresponds to an intuitive distinction between thought and language that has no analog in the narrow/wide content split. Furthermore, just as one can give up the expressive constraint without distinguishing narrow from wide content, one can posit a narrow/wide content split while accepting the expressive constraint. For instance, McGinn 1982 relies heavily on the narrow/wide distinction while explicitly endorsing the expressive constraint.

conventional meaning are clearly distinguished, giving an account of thought and utterance content that accommodates the constraint may be easier than supposed.

While the expressive constraint posits an ‘internal’ connection between thought and utterance content, it involves no claim to priority on the part of either thought or language.<sup>18</sup> Consequently, the two most obvious ways of accommodating the expressive constraint involve understanding one of either thought or utterance content in terms of the other. We can call accounts that take thought contents as basic (and try to understand utterance content entirely in terms of the thoughts that our words express) “belief-theoretic accounts of meaning.” Such accounts focus primarily on the role of language in thought, tend to be individualistic, and are tied to fairly robust conceptions of self-knowledge, language mastery and behavioral explanation. On the other hand, we can call accounts that take linguistic meaning in a public language as basic (and try to understand thought contents entirely in terms of the meaning of the sentences we use to express them) “meaning-theoretic accounts of belief.”<sup>19</sup> Such accounts focus primarily on the role of language in communication, tend to be non-individualistic, and are tied more closely to issues relating to truth and assertion. Belief-theoretic accounts have the most intuitive pull when we focus on such topics as the contents of our thoughts and our knowledge of what we mean, while meaning-theoretic accounts have the most intuitive pull when we focus on such topics as our ability to communicate, the content of what we say, and the possibility of error.

Both belief- and meaning-theoretic accounts are reductive. Belief-theoretic accounts are committed to explaining what determines the content of the speaker’s beliefs without helping

---

<sup>18</sup> For instance, the contrasting positions of Chisholm and Sellars (in their famous “Correspondence on Intentionality” (Sellars and Chisholm 1958)) both satisfy the constraint. <sup>19</sup> Of course it might seem that a meaning theoretic account could take the *individual’s* language as primary, and then derive beliefs from that. I will not discuss this possibility here, though (as will be discussed in the next section) theories which ostensibly have such a goal usually turn out to be belief-theoretic, since, to account for jokes, lies, etc, the individual’s language is usually taken simply to be *evidence* for the thoughts behind it.



itself to the (conventional) meanings of the words he uses. Meaning-theoretic accounts are committed to explaining what determines the meaning of a speaker's words without helping itself to his beliefs. Each requires that one explain either the content of the speaker's thoughts or the content of his utterances without appealing to the other. However, the resulting 'purity' of such theories leaves them with a number of unintuitive consequences. In particular, belief-theoretic accounts drive too large a wedge between conventional meaning and belief content, while meaning-theoretic accounts often fail to account for the possibility of their being such a wedge at all.

Philosophical accounts of thought and language have traditionally been belief-theoretic.<sup>20</sup> However, if belief contents are to be assigned independently of any reference to what a speaker's words mean in a public language, then both belief and utterance contents should be individualistic. This presents a problem for belief-theoretic accounts because the work of Kripke, Putnam and Burge has undermined such semantic individualism, and thus brought into question whether an adequate account of either belief or utterance content could be given within a belief-theoretic framework.<sup>21</sup> For instance, my own beliefs and linguistic capacities often underdetermine what I am intuitively taken to refer to by my terms. (To return to an earlier example, I'm taken to refer to beeches by "beech" and elms by "elm" even if I can't tell the two apart.) Consequently, belief theoretic accounts will have to treat our thought contents as being far less determinate than typically assumed. All of the reasons that might favor the assignment of *elm* to my term "elm" involve reference to the meaning *conventionally* associated with "elm," and this is precisely the sort of information that the belief-theorist cannot allow to be relevant.<sup>22</sup>

Furthermore, if conventional meanings do not affect utterance content, our learning what another believes from his sincere utterances can seem problematic. The content of our utterances may reflect the content of our beliefs, but there is no reason to think that any two speakers will

---

<sup>20</sup> See, for instance, Hobbes, 1651, Locke 1689, Schiffer 1972, Fodor, 1975, Bennett 1976, Appiah 1985, Sperber and Wilson 1986, Grice 1989, Bilgrami 1992, and Searle 1983, 1992.

<sup>21</sup> See Kripke 1972, Putnam 1975, and Burge 1979.

<sup>22</sup> Of course this is a bullet that many individualists are willing to bite. The pros and cons of individualism will not be discussed here, though the view is criticized in Jackman 1996, 1998.

mean the same thing by their words (though one may expect the two meanings to be similar). As a result, one's interlocutors will often not fully understand one's utterances, and thus they should not be expected to grasp the content of one's beliefs.<sup>23</sup>

With the 'linguistic turn' in philosophy, and the resulting shift of focus from the content of our ideas to the meanings of our words, one might think that the meaning-theoretic approach would be characteristic of many contemporary accounts of thought and utterance content. Nevertheless, while Sellarsians, Tyler Burge and Kripke's Wittgenstein are frequently treated as being in the meaning-theoretic camp, the view has been comparatively unpopular.<sup>24</sup> One can emphasize words over ideas and still privilege thought over utterance content provided that one understands thought itself as being essentially linguistic. The priority of language does not ensure a priority of *utterance* content.

The unpopularity of meaning-theoretic accounts is not surprising, since such accounts bring with them a number of serious problems.<sup>25</sup> While belief-theoretic accounts are individualistic, meaning-theoretic accounts go to the opposite extreme of being *anti-individualistic*. That is to say, it's not just that the contents of the speaker's words and thoughts can be determined by factors that extend beyond his own internal states, subjective experiences and behavior (this type of position would be *non-individualistic*), but that such facts about the particular individual need play *no* role in determining what his words mean.<sup>26</sup> As a result, meaning-theoretic accounts often leave little (if any) room for a gap between the content of one's beliefs and the conventional meanings associated with the sentences one uses to express them. That is to say, no

---

<sup>23</sup> A related problem will be, of course, that many of the belief-ascriptions we make about others will be false. When the doctor claims that Bert believes (falsely) that he has arthritis in his thigh, his ascription is false because Bert believes (correctly) that he has *tharthritis* in his thigh. Indeed, Bert himself will be mistaken about this as well if he goes on to attribute beliefs about *arthritis* to his past self. (For a more extended discussion of this, see Jackman 1996.)

<sup>24</sup> See Sellars 1963, Brandom 1994, Burge 1979 and Kripke 1982. It should, however, be noted that the claim that their views are meaning-theoretic does not strike me as an entirely fair reading of any of the authors involved.

<sup>25</sup> This is to say nothing of the 'unintuitive' character of the order of explanation involved. The thought that 'meaning' belongs intrinsically to our thoughts and comes only derivatively to our words (which, on their own, are just 'lifeless sounds or marks') can seem almost inescapable at times. (This intuition is stressed heavily in Searle 1992.)

<sup>26</sup> The "non-individualistic/anti-individualistic" terminology follows Katarzyna Paprzycka's discussion of a similar taxonomy of behavioral explanations (Paprzycka 1998). Though I'm not sure if she would quite endorse my usage of the terms.

sense can be made of the potential gap that Dummett and Evans, plausibly enough, take there to be between the conventional meaning of our words and the contents of our thoughts.

It should not be surprising, then, that the further someone strays from the conventional usage of their terms, the harder it becomes to take a meaning-theoretic line towards their thoughts and utterances. We may be willing to say that Bert believes that he has *arthritis* in his thigh, but this is partly because his conception of arthritis corresponds quite closely to the conventional one. We would, however, be less willing to take such a line for someone who believed that, say, arthritis was a type of storm pattern associated with cold climates, and who would claim “it looks like the arthritis is acting up” whenever a blizzard seemed to be on its way. In such cases we may be inclined to reject the meaning-theoretic approach and say that such a speaker was talking/thinking about, say, severe winter storms when he used the word “arthritis.” Bert’s conception of arthritis is enough like ours for us to understand him as being wrong about the same disease that the doctor is right about. However, the errors envisaged about arthritis in the second case are *so* radical and far-reaching that it becomes difficult to think of the speaker as thinking about arthritis at all. In much the same way, if one invariably assigns belief contents based upon the conventional meanings of the words uttered, then one ends up with extremely unintuitive belief assignments whenever one encounters slips of the tongue, spoonerisms, or malapropisms.

Still, one thing meaning-theoretic accounts seem to explain very well is our capacity to learn what others believe from what they say. Meaning theoretic accounts not only tie belief and thought contents together, but they also treat these contents as shared by the entire community. As a result, the same content can be believed by a speaker, expressed by that speaker and grasped by her interlocutor with little problem. Meaning-theoretic accounts thus allow for a very powerful notion of linguistic expression and communication. Indeed, the problem that meaning-theoretic accounts have in this area is that they can’t make sense of our *not* coming to know what someone believes from her sincere utterances. If one does not recognize a malapropism for what

it is, one should fail to learn what the speaker believes from her utterance, but the meaning-theoretic position seems to leave no room for this sort of failure of communication.

After considering these two attempts to accommodate it, the expressive constraint might appear difficult, if not impossible, to incorporate into any picture of the relation between thought and utterance content. Belief theoretic accounts not only leave our concepts and the contents of our utterances more anemic and idiosyncratic than commonly supposed, but they also leave the idea of communication problematic. Meaning theoretic accounts seem unable to account for any sort of failure of communication within a language group, and are often forced to endorse extremely unintuitive belief ascriptions when speakers use words in non-standard ways. If these were the only ways to accommodate the constraint, then, it might seem better to give it up and provide independent accounts of each type of content. However, such a step seems unnecessary, since there is a ‘synthetic’ account of the relation between thought and utterance content that can accommodate the constraints while avoiding the problems of its reductive alternatives.

#### **4. A synthetic position**

While the belief- and meaning-theoretic accounts may be the two most obvious ways to accommodate the expressive constraint, a non-reductive account that stressed the *interdependence* of belief content and linguistic meaning could, if available, capture the virtues, and avoid the faults, of its two reductive rivals. Indeed, it will be argued here that such reductive accounts can be avoided by properly appreciating that we have beliefs not only about the world around us, but also about the language with which we think about this world. Conventional meanings enter into the contents of our thoughts because we often have beliefs that commit us to being understood in terms of those conventional meanings. One can find a middle ground between theories that tie content exclusively to belief and those that tie it directly to conventional meanings by recognizing that one cannot take seriously all of the former without including the subject’s relation to the latter.

‘Interpretational theories’ (that is, theories that give pride of place to the process of “radical” interpretation) might seem well placed to find such a middle ground. Indeed, an account of the relation between thought and language that stressed such interdependence would seem in keeping with Davidson’s claim that “Neither language nor thinking can be fully explained in terms of the other, and neither has conceptual priority.”<sup>27</sup> Nevertheless, interpretational theories can slide remarkably easily into a belief-theoretic form.<sup>28</sup> The theories of meaning and belief are supposed to be “interrelated constructs of a single theory,”<sup>29</sup> but the major constraint on the joint theory, the Principle of Charity, has more to do with what the speaker can be taken to believe than with what he can be taken to say.<sup>30</sup> Indeed, philosophers in this tradition have frequently been quite hostile to the suggestion that conventional meanings have *any* effect on what a speaker means by his words. Quine, for instance, refuses to treat what the interpretee’s compatriots have to say as being at all relevant to how his words should be translated, and he treats the interpretee’s linguistic interactions with other members of his community as “contamination” of the interpreter’s evidence.<sup>31</sup> In much the same way, Davidson is reluctant to see social usage as having anything to do with what we mean, and Bilgrami treats endorsing such contributions as incompatible with our conceptions of self-knowledge and rationality.<sup>32</sup> Consequently, while interpretational accounts might be developed in ways that are not belief-theoretic, the most prominent writers working within this tradition have tended to understand interpretation this way.

Nevertheless, there is reason to think that interpretational theories need not (indeed, should not) slide towards the belief-theoretic style of account favored by Quine, Davidson, and Bilgrami. When one interprets someone, one wants to capture their perspective on the world

---

<sup>27</sup> Davidson, 1975, p. 156.

<sup>28</sup> This tendency is most explicit and unapologetic in Bilgrami 1992.

<sup>29</sup> Davidson 1974, p. 146. <sup>30</sup> This is implicit in Wilson’s original discussion of the Principle (see Wilson 1959, p. 531), and this shifting of Charity from the domain of language to that of belief becomes increasingly explicit in Quine and Davidson’s work.

<sup>31</sup> Quine 1960, pp. 37-8.

<sup>32</sup> Davidson 1982, 1986, 1992. Bilgrami 1992.

while not getting so caught up in it that one is unable to see where their perspective is distorted. A good interpretation should enable one to see *one's own* world through *the interpretee's* eyes. Since capturing the *interpretee's* perspective on the world is the goal of the interpretation, it is natural that her self-interpretations should have a certain constitutive force for it. We try to make our interpretation of her match what her self-interpretation would be if she were aware of all that we were.<sup>33</sup> This requires that we try to understand the speaker as she would understand herself, and it is a synthetic position, rather than a belief-theoretic one, that is best able to do this.

Someone adopting a synthetic position should (like the belief-theoretician) stress the importance of capturing the interpretee's point of view, but (unlike the belief-theoretician) also recognize the importance of the speaker's implicit assumptions about the public language and her relation to it. If this is done, one can allow the meanings conventionally associated with the speaker's words to play a large role in determining what she means, without losing hold of the idea that an interpretation should capture her perspective on the world. Belief-theoretic accounts of interpretation get much of their motivation from a perceived need on the interpreter's part to capture the interpretee's point of view. However, an integral part of a speaker's point of view is her assumption that she shares a language with her fellows. The speaker need not consciously formulate the belief that she speaks the same language as her fellows. Rather, this assumption is manifested in behavior such as her attitude ascriptions (such as willingness to move from "John said 'I have arthritis'" to "John said that he had arthritis"), her deference to correction, her willingness to admit that she doesn't know whether a particular ache is arthritis or not, or that some things that she judges to be arthritis could turn out not to be, etc.<sup>34</sup>

If the speaker's usage strays from the social norm that she believes herself to be in accord with, then (as when we are faced with any other conflict between her commitments) capturing her point of view should lead us to favor whichever element the speaker herself would favor if

---

<sup>33</sup> It is in this sense that we assume that the speaker knows what she means by her words (see Davidson 1984b, 1987). This does not, however, require that we assume that the speaker has complete mastery over how her words are correctly used. The importance of this notion of self-interpretation is discussed in greater detail in Jackman (forthcoming).

<sup>34</sup> These topics are discussed further in Jackman 1996, 1998, 2000, forthcoming.

the conflict were made manifest to her. If the speaker is unwilling to defer to the accepted usage, then we should attribute to her the false belief that she meant the same thing as her peers did by the word in question (along with a number mistaken attitude ascriptions and false beliefs about what people were talking about). If she is inclined to defer, we should treat those applications of the term that are out of line with social usage as misapplications. Since we are trying to capture the speaker's self-interpretation, this latter attribution is contingent upon the deference being understood by the speaker *herself* as a response to her own misapplication of the term. For instance, if she sees herself as deferring for purely 'pragmatic' reasons (that is to say, she understands herself as having meant something idiosyncratic by a particular word, and as changing what she means by it in order to communicate more easily), capturing her self-interpretation would not involve assigning the 'standard' referents to her past usage of her terms. As a result, if people really did defer exclusively for pragmatic reasons,<sup>35</sup> deference behavior would have little philosophical import. Still, while there certainly are cases where we alter our usage for pragmatic reasons, this certainly isn't *always* the case. People frequently defer to accepted usage because they understand themselves as having had false beliefs about which objects actually are denoted by their terms. If we take the speaker's self-interpretation to be authoritative in these matters, then we should accept such a stance on the speaker's part at face value. Consequently, if speakers defer to social usage because they take themselves to be mistaken about how their own words are properly used, then capturing their point of view will involve interpreting them as applying socially shared rather than idiosyncratic concepts.

The strong *anti*-individualistic position associated with meaning-theoretic accounts, which hold that what we mean by our terms is *always* determined by social usage, receives no support from this argument. There are clearly cases where we diverge from social usage and, when this divergence is pointed out, see ourselves as mistaken about social usage rather than the objects in question. This is why we can ascribe socially determined concepts like *arthritis* to Bert, but cannot do so to the person who thinks that 'arthritis' is a type of weather pattern. Bert will tend

---

<sup>35</sup> Davidson and Bilgrami have argued that they should be understood this way (Bilgrami 1992, Davidson 1994), and for a critical discussion of their arguments, see Jackman 1996, 1998.

to view himself as having a false belief about arthritis, while the latter speaker typically won't view himself this way. Rather, he will understand himself as having expressed a belief about the weather in a very idiosyncratic way. However, a *non-individualistic* position, which requires only that *some* of our concepts *can* be tied to social usage, does follow from the synthetic position outlined above.

This synthetic position can thus accept the type of non-individualist content ascriptions Davidson rejects while holding on to his insight that the tie between the meaning of a speaker's utterance and what others mean by the same words is "neither essential nor direct," but rather "comes into play only when the speaker intends to be interpreted as (certain) others would be."<sup>36</sup> The motivation behind the theory of interpretation can thus be understood as leading to this non-individualistic synthetic position, rather than the individualistic one associated with belief-theoretic accounts.

The synthetic position has the advantages of both the belief and the meaning theoretic views. On the one hand, it allows us to capture the agent's point of view just as well as, indeed better than, belief-theoretic accounts. On the other hand, it accounts for our intuitions about how what we mean by our terms can go beyond our current recognitional capacities, and provides a good account of how we come to know what others believe from their sincere utterances. While the synthetic account provides room for a distinction between what a person means by his words and the conventional meanings associated with those words, it also explains why these will typically be the same. We generally intend to be interpreted as our fellows are, and while such an intention is defeasible (if, for instance, we make a slip of the tongue or our usage strays too far from the public norm), minor differences between ours and public usage usually will not be enough to make us give it up. While we occasionally may not assent to the attribution of standard meanings to our words, these cases tend to be exceptional, and we usually consent to be

---

<sup>36</sup> Davidson 1992, 261. It is thus not always helpful to discuss issues in this area in terms of 'individualism' at all, since the 'non-individualistic' position outlined above is still *methodologically* individualistic, that is, the relevance of the 'external' factors are ultimately justified by the individual's own attitudes towards them. (For a discussion of this, see Jackman 1998.) In this respect my usage may be at odds with Paprzycka 1998.



interpreted (and, crucially, interpret ourselves) in accordance with the conventional meanings.<sup>37</sup> While the speaker and hearer may have different beliefs associated with the terms used in a particular utterance, both will usually be willing to defer to conventional usage. As a result, both can be taken to mean the same thing by the sentence, and to have thoughts with the corresponding content.

The expressive constraint can thus be accommodated within a synthetic account of the relation between thought contents, utterance contents, and the contents conventionally associated with our words. The sort of interpretational picture suggested above avoids separate explanations of thought and utterance content without reducing one to the other. As a result, it explains both why these two types of content will correspond with each other, and why they will often (but not always) be tied to conventional meaning. Furthermore, the proposed explanation avoids the pitfalls associated with the belief- and meaning-theoretic accounts. Both belief- and meaning-theoretic accounts presuppose that conventional meaning can only be *relevant* to what our own words mean by being *equated* with it, and once this assumption is given up, one can give a much more plausible account of the relation between thought and utterance content. The tendency to view the expressive constraint as requiring either a belief- or meaning-theoretic approach is a large part of the reason why it can seem so hard to give an account of thought and utterance content that accommodates it. Consequently, the possibility of such a synthetic account makes the prospects for accommodating such an internal connection between thought and utterance content seem much more promising.

---

<sup>37</sup> Indeed, the fact that Bert is speaking to a *doctor* and not, say, his brother may be relevant here. When we are actually consulting an expert, the presumption (on the part of both participants) that we will intend to be following ‘expert usage’ will be considerably stronger than when we are talking within a possibly idiosyncratic sub-community such as our family. This is discussed in greater detail in Jackman 1996, 1998, 2000.

## BIBLIOGRAPHY

- Appiah, A. 1985. *Assertion and Conditionals*, London: CUP.
- Bennett, J. 1976. *Linguistic Behavior*, New York: CUP.
- Bilgrami, A. 1992. *Belief and Meaning*, Cambridge: Blackwell.
- Block, N. 1986. "Advertisement for a Semantics for Psychology", in French, Uehling & Wettstein. 1986. *Midwest Studies in Philosophy X: Studies in the Philosophy of Mind*, Minneapolis, University of Minnesota Press.
- Brandom, R. 1994. *Making it Explicit*, Cambridge: Harvard University Press.
- Burge, T. 1979. "Individualism and the Mental", in French, Uehling & Wettstein. 1979. *Midwest Studies in Philosophy IV: Studies in Metaphysics*, Minneapolis, University of Minnesota Press.
- Burge, T. 1989. "Wherein is Language Social?", in George, A. (ed.) *Reflections on Chomsky*, New York: Blackwell, 1989.
- Davidson, D. 1974. "Belief and the Basis of Meaning", in Davidson 1984a.
- Davidson, D. 1975. "Thought and Talk", in Davidson 1984a.
- Davidson, D. 1982. "Communication and Convention", in Davidson 1984a.
- Davidson, D. 1984a. *Inquiries into Truth and Interpretation*, Oxford: OUP.
- Davidson, D. 1984b. "First Person Authority" *Dialectica*, Vol. 38 no, 2-3.
- Davidson, D. 1986. "A Nice Derangement of Epitaphs", in LePore, E, (ed.). *Truth and Interpretation*, Oxford: Blackwell, 1986.
- Davidson, D. 1987. "Knowing One's Own Mind", *Proceedings and Addresses of the American Philosophical Association*.
- Davidson, D. 1992. "The Second Person", in French Uehling and Wettstein. 1992. *Midwest Studies in Philosophy XVII: The Wittgenstein Legacy*, Notre Dame, University of Notre Dame Press.
- Davidson, D. 1994. "The Social Aspect of Language", in McGuinness and Oliveri (eds.). 1994. *The Philosophy of Michael Dummett*, Dordrecht: Kluwer.
- Dreyfus, H. 1991. *Being in the World*, Cambridge: MIT press.
- Dummett, M. 1991. *The Logical Basis of Metaphysics*, Cambridge: Harvard University Press.
- Dummett, M. 1993. *Origins of Analytic Philosophy*, Cambridge: Harvard University Press.
- Evans, G. 1982. *The Varieties of Reference*, New York: OUP.
- Field, H. 1978. "Mental Representation", in Block, N. (ed.). *Readings in the Philosophy of Psychology, volume 2*, Cambridge: MIT, 1981.

- Fodor, J. 1975. *The Language of Thought*, New York: Crowell.
- Grice, P. 1989. *Studies in the Way of Words*, Cambridge: Harvard University Press.
- Hobbes, T. 1651, 1981. *Leviathan*, New York: Penguin.
- Jackman, H. 1996. *Semantic Norms and Temporal Externalism*. Ph.D. Thesis, University of Pittsburgh.
- Jackman, H. 1998. "Individualism and Interpretation" *Southwest Philosophy Review* V.14, n. 1, Jan 1998.
- Jackman, H. 2000. "Deference and Self-Knowledge" *Southwest Philosophy Review*, V.16, n. 1, Jan 2000.
- Jackman, H. (Forthcoming), "Charity, Self-Interpretation, and Belief." *Journal of Philosophical Research*. Forthcoming.
- Kripke, S. 1972. *Naming and Necessity*, Cambridge: Harvard University Press.
- Kripke, S. 1979. "A Puzzle about Belief", in Salmon, N. and Soames, S. (eds.) *Propositions and Attitudes*, New York: OUP, 1988.
- Kripke, S. 1982. *Wittgenstein on Rules and Private Language*, Cambridge: Harvard University Press.
- Lewis, D. 1979. "Attitudes *De Dicto* and *De Se*". In Lewis, D. *Philosophical Papers, Volume 1*. New York: OUP. 1983.
- Loar, B. 1985. "Social Content and Psychological Content", in Grimm, R & Merrill, D. *Contents of Thought*, Tuscon: University of Arizona Press, 1988.
- Locke, J. 1689, 1975. *An Essay Concerning Human Understanding*, New York: OUP.
- McGinn, C. 1982. "The Structure of Content", in Woodfield, A. (ed.) *Thought and Object*, New York: OUP, 1982.
- Moore, G. E. 1993. *Selected Writings*. New York: Routledge.
- Paprzycka, K. 1998. "Collectivism on the Horizon: A challenge to Pettite's Critique of Collectivism" *Australasian Journal of Philosophy*, vol. 76, no. 2, pp. 165-181.
- Putnam, H. 1975. "The Meaning of 'Meaning'", in Putnam, H. *Mind Language and Reality*, New York: CUP, 1975.
- Putnam, H. 1981. "Reference and Understanding", in Putnam, H. *Meaning and the Moral Sciences*, Boston: Routledge & Kegan Paul, 1981.
- Quine, W.V.O. 1960. *Word and Object*, Cambridge: MIT.
- Schiffer, S. 1972. *Meaning*, Oxford: OUP.
- Searle, J. 1983. *Intentionality*, New York: CUP.

Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge: MIT.

Sellars, W. 1963. *Science Perception and Reality*, London, RKP.

Sellars, W, and Chisholm, R. 1958. "Correspondence on Intentionality", in Feigl, H. Scriven, M. and Maxwell, G. (eds.) *Minnesota Studies in the Philosophy of Science II: Concepts, Theories and the Mind-Body Problem*. Minneapolis, University of Minnesota Press, 1958.

Sperber & Wilson. 1986. *Relevance*, Cambridge, Harvard UP.

Wilson, N. L. 1959. "Substance without Substrata", *Review of Metaphysics*, 12 (1959), pp.521-539.

## **Expression, Thought, and Language**

### **ABSTRACT**

This paper discusses an “expressive constraint” on accounts of thought and language which requires that *when a speaker expresses a belief by sincerely uttering a sentence, the utterance and the belief have the same content*. It will be argued that this constraint should be viewed as expressing a conceptual connection between thought and language rather than a mere empirical generalization about the two. However, the most obvious accounts of the relation between thought and language compatible with the constraint (giving an independent account of one of either linguistic meaning or thought content and understanding the other in terms of it) both face serious difficulties. Because of this, the following will suggest an alternative picture of the relation between thought and language that remains compatible with the constraint.