THESIS

DONALD DAVIDSON: MEANING, TRIANGULATION,

AND CONVENTION

Submitted by

Jesse A. Thompson

Department of Philosophy

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Summer 2011

Master's Committee:

     Advisor: Michael Losonsky

     Katie McShane
     Paul Trembath

ABSTRACT

DONALD DAVIDSON: MEANING, TRIANGULATION,

AND CONVENTION

This thesis examines a narrow portion of Donald Davidson's work in the philosophy of language, specifically his theory of utterance meaning put forward in the essay "The Second Person". In light of certain counterexamples I first attempt to adjust Davidson's theory, guided by a comment that suggests Davidson's awareness of the problem. When no satisfactory means of amending Davidson's proposal is found, I turn to the alternative proposals rejected by Davidson in the hopes of finding motivation to continue pursuing Davidson's arguments.

The second chapter is devoted to Davidson's rejection of what I call the "subjectivist" position. I contend that while Davidson provides a strong argument against subjectivism, that argument entails further complications that Davidson fails to resolve. Since an adequate rejection of a position should involve an alternative that reduces or eliminates difficulties rather than simply transforms them, I conclude that Davidson has failed to motivate his move away from subjectivism.

The third chapter is a discussion of Davidson's arguments against what I call "conventionalism". Here I show that while Davidson argues convincingly against a particular role that conventions might play, he does not motivate a move away from a broader understanding of conventionalism. At best, Davidson's arguments show that the conventionalist position should be amended, but not that it should be rejected.

The result is that Davidson's theory enjoys a rather deflated place among theories of utterance meaning. The alternatives, while suffering from their own defects, are no more problematic than Davidson's theory. Consequently, we ought to consider each of these theories as possible solutions to the problem of utterance meaning.

TABLE OF CONTENTS

`

Chapter One

How is it possible that when a speaker stands before a hearer and emits an acoustic blast such remarkable things occur as the speaker means something; the sound he emits mean something; the hearer understands what is meant; the speaker makes a statement, asks a question, or gives an order? How is it possible, for example, that when I say "Jones went home", which after all is in one way just a string of noises, what I mean is: Jones went home.

-John Searle, *Speech Acts*

As Searle observes in the first paragraph of "Speech Acts," human communication is a mysterious thing. At the same time, however, there is no doubt that we often do manage to understand one another, and that the noises we make usually have some meaning. But on what basis do we distinguish between a speaker's meaningful utterances and mere noise-making? In order to begin a proper examination of the efforts to answer this question, we must first distinguish this question from two others. The question being asked here is a matter of the actual meaning of particular utterances made by particular speakers at particular times. We will call this utterance meaning.

By contrast, we are not concerned with what we will call abstract meaning. When we talk about language itself, we speak as if words possess some meaning independent of their particular utterance or use. For example, philosophers like Saul Kripke consider this sort of meaning to be determined by the appropriate "assertability conditions" for every possible utterance in a given language, where assertability

conditions are the conditions (or contexts) under which it is acceptable to assert a proposition. For Donald Davidson, abstract meaning is determined by the truth conditions that can be specified for every possible utterance in a given language, where truth conditions are the conditions under which a proposition is true.[1] Assertability and truth conditions are abstract in the sense that they do not determine the meaning of any actual utterance, but only the meaning of utterances removed, or abstracted, from actual instances of use.[2] This sort of talk about language is useful as a theoretical tool, but by itself does little to explain how a speaker's actual utterances possess meaning and are successfully interpreted by others as having that meaning. Davidson's remarks on language as an abstract object are similar, saying "we want to understand the actual utterances of others, and we want our utterances to be understood. What has language to do with this?"[3] Knowing the appropriate assertability or truth conditions for a language does not tell us what a particular speaker's utterance means on a particular occasion. We are interested in the conditions that allow a speaker's utterance to have meaning for others, and it is for this reason that we will focus on utterance meaning.

There is another variety of meaning to be distinguished from utterance meaning which we will call speaker meaning. Speaker meaning is the content that a speaker has in mind when making a particular utterance. Though utterance meaning is a property of particular utterances made by particular speakers, it may be distinct from the meaning that the speaker intends to communicate. The speaker meaning of an utterance, then, is simply what a speaker intends to communicate, independent of the utterance he makes.

---

[1] Donald Davidson, "Truth and Meaning," *Synthese*, 17, (1967).
[2] This is not to say that abstract meaning has no relation to actual use, simply that the meaning of speaker's actual utterances (the utterance meaning of an utterance) can be distinct from abstract meaning.
[3] Donald Davidson, "The Second Person," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 109.

In Davidson's words, the distinction between speaker meaning and utterance meaning is the distinction "between what a speaker, on a given occasion, means, and what his words mean."[4] We are concerned with what words, as used by particular speakers at particular times, mean. Perhaps it is clarifying to add that often in instances of miscommunication, utterance meaning fails to coincide with the speaker meaning of that utterance.[5]

Since these distinctions are sometimes difficult to see, some examples are in order. The distinction between speaker meaning and abstract meaning is most clear in instances of communication involving a second language. Suppose I have been taught that "Tengo un plátano" conventionally translates as "I am sorry" in Spanish, when in fact it conventionally translates as "I have a banana." On a trip to Mexico, I accidently bump into a person, and say "Tengo un plátano," intending that my utterance be interpreted as meaning "I'm sorry." Clearly the speaker meaning of my utterance is "I'm sorry," since that is the meaning I intend my utterance to have. It is just as clear that the abstract, or conventional, meaning of my utterance is "I have a banana," since that is what those words abstracted from use, say in a dictionary, mean. Conventional meaning and speaker meaning clearly diverge in this case.

The distinction between abstract meaning and utterance meaning is clearest in cases of malapropisms. In his essay "A Nice Derangement of Epitaphs," Davidson provides several examples that illustrate this distinction.[6] If I say that "Peyton Manning is the pinochle of success," it is at best unclear what the abstract, or conventional,

---

[4] Donald Davidson, "A Nice Derangement of Epitaphs," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press), 252.

[5] What we are calling "utterance meaning" is what, for Davidson, is arrived at once successful interpretation has occurred.

[6] Donald Davidson, "A Nice Derangement of Epitaphs," in *The Essential Davidson*, (New York: Oxford University Press, 2006).

meaning of that utterance would be.[7] However, even in the absence of a defined context, interpreters can easily recognize that I intend to be interpreted as saying that Peyton Manning is the pinnacle of success. Davidson explains, "the hearer realizes that the 'standard' interpretation cannot be the intended interpretation…The absurdity or inappropriateness of what the speaker would have meant had his words been taken in the 'standard' way alerts the hearer to trickery or error; the similarity in sound tips him off to the right interpretation."[8] The abstract meaning of the utterance is absent or nonsensical, yet on this occasion my utterance conveys a meaning that is understood by my interpreters as I intended. In cases like this, abstract meaning and utterance meaning are clearly distinct.

The distinction between utterance meaning and speaker meaning is plain in cases of simple misspeaking, instances in which a speaker makes an utterance with the intention of conveying a certain meaning, but accidentally uses words that do not successfully convey that meaning. For instance, after talking about a bicycle I want to purchase, I intend to say to my friend, "would you get me the chips?" but instead say "would you get me the bicycle?" In my head, the content was, so to speak, "would you get me the chips?" but that was not the meaning expressed by my utterance.   Neither did I intend that "would you get me the bicycle?" should be interpreted as "would you get me the chips?" In this case, my words did not express what I intend to express, so the utterance meaning differs from the speaker meaning of my utterance.

---

[7]The cause of the mistake is not important. Perhaps I am confused about the conventional meaning of the word "pinochle", perhaps it was a slip of the tongue, or perhaps I am being devious.

[8]Donald Davidson, "A Nice Derangement of Epitaphs," in *The Essential Davidson*, (New York: Oxford University Press, 2006), 252.

We may now ask specifically "what constitutes utterance meaning?" as opposed to what constitutes abstract meaning or speaker meaning.   That is, what must be the case for our actual utterances to have meaning? It must also be clear that the answer to this question is explanatory, not prescriptive. There is little doubt that some of our utterances have meaning and others do not, and we do not often disagree about which utterances have meaning and which do not. The matter of debate regards the explanation of that phenomenon. The answer should give us insight into the concept of meaning in general, and it is worth examining some previously proposed answers to this question. The first proposal emerged in the Enlightenment, and places utterance meaning entirely in the mind of the speaker. The second proposal, often attributed to Wittgenstein, places utterance meaning entirely in a public, social arena. The third proposal, with which this thesis is most concerned, is Donald Davidson's and incorporates both subjective and social elements, though in different ways than its predecessors.

Utterance meaning and the mental are clearly closely linked, and the Enlightenment empiricists relied heavily on mental content to deliver their account of utterance meaning. For empiricists like Locke, there is an external world, independent of mind, and then there are human minds that experience the external world mediated through subjective, private, infallible ideas, or sense data. Under this theory, words express thoughts, and the particular thoughts that words express entirely determine their meaning; what my utterances mean is a matter of what thoughts I choose to express through my utterances. Whether or not my audience manages to interpret my utterances as expressing those thoughts does not affect the meaning of my utterances.

Just as we have privileged access to our sense data, we have privileged access to what we know, believe, intend, and all the other contents our mind. This access then

extends to knowledge of the utterance meaning of the words we use to express those very things that we perceive, believe, and intend. Since according to the empiricists, one cannot be mistaken about any of the contents of one's mind, neither can one be mistaken about the meaning of one's words. In other words, speaker meaning determines utterance meaning. If we state this standard of utterance meaning more formally, it is clear that it can be met in absence of any interpretations or understanding of others:

S's utterance **U** means that **P** only if **S** expresses the thought that is **P** by uttering **U**.

The Enlightenment, then, had a highly individualistic standard of utterance meaning; what an utterance means is entirely determined by the private contents of the speaker's mind. In the middle of the 20th century, Wittgenstein argued for a more social conception of utterance meaning. Wittgenstein's basic argument is this: to speak meaningfully is to follow a particular rule or set of rules, and rule-following is not something that can be done in isolation. The distinction between *actually* following a rule and merely *thinking* that one is following a rule (without actually following it) can only be made by comparing that individual's behavior to the behavior of a rule-following community. Therefore, in order for one's utterances to be meaningful in virtue of being in accordance with a rule, one must speak as others speak.[9] Shared rule-following of this sort is a convention, and we may therefore characterize the second standard of utterance meaning as follows:

---

[9]This requirement is obviously not limited to language. According to this view, one cannot be said to follow any rule of any sort in the absence of a comparable rule-following community.

S's utterance **U** means that **P** if and only if there exists a convention that utterances of **U** mean that **P.**[10]

Donald Davidson finds both proposals inadequate. Davidson rejects the first standard due to his rejection of subjective states as traditionally conceived. For Davidson, subjective states do not require the existence of private mental objects with which only the subject is acquainted. According to Davidson, we are causally related to the world in certain ways, and the manner of those causal relations *in part* determines the things we believe, intend, and mean. According to the empiricists, objects of thought that could be described individualistically would have to be objects that "'Are what they seem and seem what they are' – that is, have all and only the properties we think they have," for if objects could have properties which Alas, for Davidson, "there are no such objects."[11] The private, infallible objects of thought, which the Enlightenment empiricists affirmed and made the basis of their standard of utterance meaning, do not exist for Davidson, and as a result he finds that their standard of meaning is lacking.

The second standard Davidson rejects not because it is social, but because it claims that utterance meaning is fundamentally conventional. For the conventionalist, to speak a language meaningfully is to follow conventions.[12] While Davidson does not deny that linguistic conventions exist and are often followed, he does not consider them *necessary* conditions of meaningful speech. Davidson's strongest argument against convention is a standard modus tollens: if conventions are necessary conditions of successful communication, then successful communication in the absence of

---

[10] The existence of convention implies rule-following and being shared (as well as serving a common interest and being common knowledge).

[11] Donald Davidson, "What is Present to the Mind?" in *Subjective, Intersubjective,* (New York: Oxford University Press, 2001), 54.

[12] " Convention" can be used interchangeably with "shared rule-following"

conventional speech does not exist. Successful communication in the absence of

convention *does* exist; therefore conventions are not necessary for successful

communication. By rejecting convention as a necessary condition of communication,

Davidson rejects the conventional standard of utterance meaning.

The necessary conditions of utterance meaning suggested by Davidson himself

are neither entirely individualistic nor conventional. In his essay "The Second Person,"

Davidson proposes that the "distinction between thinking one means something and

actually meaning it can be made in terms of the success of the speaker's intention to be

interpreted in a certain way."[13] The distinction Davidson is characterizing here is simply

the distinction between speaker meaning and utterance meaning. Providing a means of

distinguishing between utterances that are merely intended to be meaningful and

utterances that actually are meaningful is to provide a standard of utterance meaning.

This standard can at least initially be characterized as follows:

S's utterance **U** means that **P** if and only if **S** intends to be interpreted as meaning

P by uttering **U** and **S**'s intention is satisfied.[14]

This formulation is itself in need of some interpretation, but it is a sufficient

starting point. The concept most in need of clarification is satisfaction of an intention.

The contents of intentions are states of affairs which are believed not to be actual. The

nature of intention is such that the existence of an intention entails the belief that the

intender is capable of actualizing the content of the intention, and they have some degree

---

[13]Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 117.

[14]If it seems unsatisfactory to employ the concept of meaning itself, remember that we are not here attempting to give a definition of meaning, but merely the necessary conditions which must be met in order for utterance meaning to occur.

of expectation that they *will* actualize that content. Satisfaction of an intention should then in some sense involve the actualization of that content. If I have an intention to stand up, but find myself incapable of movement and remain sitting, I will have failed to satisfy my intention to stand up.

When communicating, we must intend our words to be interpreted as having some particular meaning. Without this requirement, there would be no way of making a distinction between instances of speech and other instances of noise-making. My inadvertent sighs, grunts, and mutterings would have to receive the same treatment as my meaningful speech, which does not fit our practice.   We regularly distinguish between meaningful speech and mere noise-making, and we do so by distinguishing between intentional and unintentional noises. Even when the intention to be interpreted is present, however, we still possess the capacity to make mistakes; we can intend for our utterances to be interpreted in a certain way even when our utterances go uninterpreted, or are even uninterpretable.[15]

Now, the initial interpretation of Davidson's proposal can be modified as follows. The satisfaction of an intention should involve the actualization of the content of that intention. So Davidson's proposal would look more like this:

S's utterance **U** means that **P** if and only if **S** intends to be interpreted as meaning that **P** by uttering **U** and **S**'s utterance **U** is *actually* interpreted as meaning that **P**

We have now replaced the notion of satisfaction with a description of the actualized intentional content. In this case, **S**'s intention becomes actual. **S** is in fact interpreted as she intended. Let us call this the "Actualized Proposal" or AP. This is the instinctive

---

[15]Of course, in order to have an intention to be interpreted in a certain way, one would have to at least *believe* that is possible for her to be interpreted as she intends, but that does not entail that it is actually possible.

interpretation of Davidson's proposal, and is textually supported as well. "The presence of intentions, "says Davidson, "gives content to an attribution of error by allowing for the possibility of a discrepancy between intention and accomplishment."[16] Four pages later, "the intention of the speaker to be interpreted in a certain way provides the 'norm'; the speaker falls short of his intention if he fails to speak in such a way as to be understood as he intended."[17] Talk of a 'discrepancy between intention and accomplishment' and 'failure to be understood as one intends' must, without significant acrobatics, be understood as an endorsement of the AP.

If Davidson is endorsing the AP, or something like it, then he must answer a set of counterexamples. It is easy to imagine a situation like the following, and it certainly happens with some frequency. I am having a conversation with my friend as we pass by a tractor-trailer waiting at a stop light. As I am about to say "Yes, I have no bananas," the light turns green, and the vehicle in front of the truck does not move. As I speak the sentence as I intended, the truck lets out a frustrated blast of its air-horn, making it impossible for my friend to make out what I have said. If Davidson truly endorses the AP, he would have to admit that although my intention to be interpreted in a certain way *could* have succeeded, and was indeed very likely to succeed, in actuality it did not succeed, and therefore my utterance was without utterance meaning.

Of course, my utterance could still possess speaker meaning and abstract meaning, but it is the essence of utterance meaning that Davidson intends to capture. I certainly intended my utterance to communicate a certain content, and there are definite truth conditions for my utterance which my friend would have assigned to my utterance.

---

[16] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 112.
[17] Ibid., 116.

If my words are void of utterance meaning, however, my utterance itself would be mere gibberish, and this is where our intuitions about our linguistic practices depart from the dictates of the AP. It seems clear that in such a case, I have done something more than make a meaningless utterance. My utterance was meaningful, my friend simply failed to grasp it due to an obstructing noise.

This counterexample can take many forms, since there are many possible ways for an utterance to go un-interpreted or to be interpreted in a way not intended by the speaker. My friend could be ignoring me, distracted, or simply not paying attention to my utterance rather than have the utterance be obstructed. The utterance could be partially, instead of fully obstructed, leading to a misinterpretation rather than a failure to interpret entirely. The salient element in all these possibilities is the existence of a factor that would affect the meaning of an utterance but is ordinarily considered to have no effect on the meaning of our words. It would be strange, if not absurd, to claim that the utterance meaning of one's utterance is somehow dependent on the attention of one's audience or the absence of an obscuring noise.

Neither can Davidson claim that in this sort of case the utterance retains utterance meaning in virtue of the fact that the speaker herself successfully interprets her own utterance, though others do not. The consequence of such a move would be that every utterance with speaker meaning would possess utterance meaning, since a speaker can always interpret her own utterances as she intends them to be interpreted. This would undermine the very distinction between speaker meaning and utterance meaning that Davidson demands we maintain. In addition to blurring the line between speaker meaning and utterance meaning, the move would threaten another tenet of Davidson's theory, the demand for a distinct interpreter. If a speaker can truly interpret their own

utterances in the same way that a distinct person interprets a speaker's utterances, there is

nothing to eliminate the possibility of a private language, which is a possibility Davidson

explicitly denies.[18]

Is the AP an accurate interpretation of Davidson's proposal? If we are charitable

to his position, we must conclude that Davidson has a more nuanced position in mind

despite the textual evidence that he endorses the AP. What other interpretations are

available? Davidson provides the beginnings of an alternative proposal in a footnote in

"The Second Person":

> There is a point here I have not accommodated. A speaker fails in an
> intention if he is not interpreted as he intends. But it would be wrong
> to say that such a failure is necessarily a failure to give the meaning to
> his words that he intended the interpreter to catch. The latter failure
> depends (in ways that ordinary usage may not definitively settle) on
> such questions as whether the speaker was justified in believing his
> interpreter could, or would, interpret him as he intended.[19]

Here Davidson appears to be attempting to account for the very sort of

counterexample detailed above. The counterexample is a case in which a speaker fails in

her intention to be interpreted in a certain way but her utterance still has utterance

meaning, and Davidson is suggesting conditions which could allow for just such a

situation. In fact, these would be necessary conditions, since the failure to produce

utterance meaning *depends* on these conditions. To say that **A** depends on **B** is to say that

B is a necessary condition for **A**. Davidson is proposing that a speaker's justified belief

that she could or would be interpreted as she intended is a necessary condition of

---

[18]Davidson's argument against the possibility of a private language is not relevant to this discussion, but it occurs in "The Second Person." in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001).

[19]Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 117.

utterance meaning. Davidson does not distinguish between a justified belief that the interpreter *could* interpret the speaker as intended and a justified belief that the interpreter *would* interpret the speaker as intended. Neither does Davidson make explicit exactly what the content of the justified belief should be. We will see, though, that different interpretations of these conditions result in very different standards, and that we should treat them as such.

First, let us address the content of the justified belief. The standard which we are attempting to construct on Davidson's behalf must allow for the case in which a speaker's intention fails but utterance meaning remains. The footnote says that utterance meaning can be retained if the speaker is justified in believing that her intention could or would succeed. In order to determine how such a standard needs to be constructed, we need to examine some things that Davidson says about intentions.

In "The Second Person", Davidson says that "intentions depend on the belief that one can do what one intends, and this requires that one believe nothing will prevent the intended action."[20] Another way to say this is that in order to possess an intention to do X, one must believe that the necessary conditions for doing X obtain. If I intend to stand up, then I must believe that a number of conditions obtain, for example that I possess the strength to resist the pull of gravity, that my body will respond appropriately to my commands, that the floor will not suddenly give way, and so on. If I do not believe that all of the necessary conditions of accomplishing my intention obtain, then I cannot be said to have that intention since I cannot be said to believe that I can do what I intend. So having intentions entails believing that the necessary conditions for doing what one

---

[20] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 112.

intends obtain. It is unclear, though, whether Davidson is demanding that a speaker be justified in believing that *if* the necessary conditions were to obtain, then their intention could or would succeed, or whether they must be justified in believing that the necessary conditions *do* obtain and that under those conditions their intentions could or would succeed. Put more formally:

> **S** is justified in believing that if conditions **C** were to obtain, **S** could or would be interpreted to mean that **P** by uttering **U**.

Or

> **S** is justified in believing that conditions **C** obtain and if **C**, **S** could or would be interpreted to mean that **P** by uttering **U**.

For Davidson, the intention to **X** entails the belief that the necessary conditions of **X**ing obtain, and the intention to be interpreted in a certain way is a necessary condition of utterance meaning, so the belief that the necessary conditions of being interpreted as one intends obtain is a necessary condition of utterance meaning. Davidson must then affirm the latter characterization and not the former, since only the latter proposal requires that the speaker possesses the belief that the necessary conditions of being interpreted as one intends actually obtain. The content of the justified belief must be that the speaker is justified in believing that the necessary conditions for being interpreted as the speaker intends obtain, and that under those conditions the speaker could or would be interpreted as she intends.

Now that the content of the justified belief has been determined, we may turn to the difference between could and would. If we modify the AP to include the justified belief that the speaker *could* be interpreted as intended, we get the following, where **C** is the necessary conditions for meaning **P** by uttering **U**:

S's utterance U means that P if and only if S intends to be interpreted as meaning that P by uttering U and S is justified in believing that conditions C obtain and that if C, S could be interpreted to mean that P by uttering U.

Let us call this the "Justificatory Proposal 1" or JP1.

In the original counterexample, even though the speaker is not actually interpreted as she intended she is entirely justified in *believing* that the necessary conditions for being interpreted as she intends have been met, and that she could be interpreted as she intends. The presence of that justified belief would be enough to satisfy the necessary conditions of utterance meaning according to the JP1 even when the intention to be interpreted in a certain way is not actualized.

It is still unclear how to interpret the JP1, however. The modal term "could" entails some sort of possibility, but what sort of possibility it is and what degree of possibility is required is not evident. Often when we speak of possibility, we intend to refer to logical possibility, that is, all possibilities which do not entail contradictions. If we interpret Davidson as implying logical possibility, we get the following:

S's utterance U means that P if and only if S intends to be interpreted as meaning that P by uttering U and S is justified in believing that conditions C obtain and that if C, it is logically possible that S's utterance U is interpreted as meaning that P

We seek a standard which distinguishes meaningful utterances from utterances lacking utterance meaning in accordance with our practices and intuitions. The above standard, let us call it the JP1, does little to provide that distinction since there would be infinitely many utterances which would possess utterance meaning but are in fact not interpretable by any existing interpreter. All that the JP1 requires is that I am justified in believing

that conditions obtain which allow for the logical possibility that I am interpreted as I intend. Given the breadth of logical possibility, it is obvious that the only conditions which I *must* justifiably believe to obtain are all but useless for our purposes. I must, say, be justified in believing that I am speaking to a being that is capable of interpretation, but that is hardly a sufficient condition for giving meaning to my words.

If I utter some made-up sounds to my friend without giving her any clues as to how I intend to be interpreted, and believing that she has no such clues, she would be right to say that I was speaking nonsense and I would have little evidence that I was not. It is hardly convincing to point out that I believe it is possible she correctly guessed at what I meant or she has mind-reading powers, thereby understanding me, no matter how justifiable it may be to believe that such things are logically possible. The JP1 would have to maintain that my utterance had utterance meaning, since by being justified in believing that the necessary conditions of being interpreted as I intend are met and that under those conditions it is logically possible that my utterance is interpreted as I intend, I have met the conditions of utterance meaning. The fact that the JP1 would consider my utterance to have utterance meaning while in practice it does not is reason enough to reject it. Since our standard must fit and explain our practice, not determine it, a standard that grants meaning to utterances which we do not consider meaningful will not suit our purposes.

While there are other ways of interpreting "possible," the ultimate difficulty with this proposal is not the modality, but the fact that the standard is composed of a justified belief. In order to see this clearly, we should briefly examine how the standard meets the current objections once the modality would is used instead of could:

16

>**S**'s utterance **U** means that **P** if and only if **S** intends to be interpreted as meaning
>
>that **P** by uttering **U** and **S** is justified in believing that conditions **C** obtain and if
>
>**C, S**'s utterance **U** would be interpreted as meaning that **P**

This standard, which we may call the JP2, seems much closer to answering the original

counterexample than the JP1, largely because instead of pointing to a merely possible

case in which the speaker's utterance is actually interpreted, it provides a conditional for

the success of the speaker's intention. Would implies a sort of possibility, but it is a more

limited sense of possibility than that implied by could. A belief that **X** would have

happened is a belief that the conditional "if ~**Y** then **X**" is true, where **Y** is the condition

or conditions which prevented **X**. In the case of the original counterexample, the

conditional would be: "if the utterance had not been obscured, then the utterance would

have been interpreted as the speaker intended." This seems to capture the intuition that

the utterance in the counterexample has utterance meaning even though it was not

interpreted as the speaker intended.

When we intend to accomplish something, we believe that we are capable of

accomplishing what we intend. When our intentions fail, we must either admit that we

were mistaken about the necessary conditions for succeeding in our intention, or that the

necessary conditions which we believed to obtain did not in fact obtain. Even if we make

these admissions, we may still possess a belief regarding the counterfactual success of

our intention. That is, we may believe that had the necessary conditions obtained, we

would have succeeded in our intention. And surely, it is a belief about counterfactual

success that in part motivates the counterexample. Once the AP is forced to deny that the

speaker's utterance in the counterexample has any utterance meaning, the intuitive

objection is something like, "He had no way of knowing that his utterance would be

17

obscured. If the truck hadn't honked, he would have been interpreted as he intended and his utterance would have possessed utterance meaning." The JP2 appears to be capable of accounting for this particular counter example.

The concept of justification certainly carries various complications, but at this point we may say that our ordinary concept of justification is sufficient for this definition, something like "a speaker's belief is justified if the evidence available to the speaker makes the belief more likely true than false". Complaints regarding the concept of justification are hardly particular to this discussion, and it is reasonable to set them aside at least for the time being.

Even if we agree to make use of the ordinary concept of justification, there is a problem with the JP2. The problem results from the fact that a justified belief can be false. In fact, in order for the JP2 to be able to avoid the counter example that plagued the AP, it *must* be possible for a justified belief to be false. If the JP2 is to respond adequately to the original counterexample, it must be possible to be justified in believing that a speaker's utterance would be interpreted as the speaker intends, even though in fact she is not. If one could only be justified in believing something that is true, then the speaker in the counterexample could not be said to have a justified belief that she would be interpreted as he intends, since it is false that he would be interpreted as he intends. One can be justified in believing something to be true which is not in fact true.

As a consequence, the JP2 is subject to an altogether new counterexample. Say I have had an especially vivid dream in which my friends and I jokingly agree that we will use the (we shall assume) previously unused verb "fleuvenate" instead of the verb "agree". The dream was so vivid that upon waking I mistake it for a memory of actual events, and return to my friends believing, and justifiably so, that if I utter the sentence "I

fleuvenate" I will be interpreted as saying that "I agree".[21] One of my friends makes an

utterance with which I agree and I, with the intention of being interpreted as meaning "I

agree", say "I fleuvenate". My friends stare at me blankly, unable to interpret my

utterance since the necessary conditions for being interpreted as I intend have not been

met, though I was justified in believing that they were, and that under those conditions I

would be interpreted as I intend. It should be clear that not only has my intention failed,

but it would not have succeeded unless my justified belief was true instead of false.

Since "fleuvenate" has never been uttered before and I have provided no clues to

my interpreters as to how I intend to be interpreted, neither my actual interpreters nor *any*

other possible interpreter could have interpreted me as I intended without some

incredible means to obtain clues as to what I meant. And yet, my belief that my utterance

would be interpreted as I intended was entirely justified, though quite false. It is also

important to note that the lack of utterance meaning results from the fact that my belief

was false, not simply from the fact that my justification was based on an inaccurate

memory. One could, of course, simply insist that the utterance in the counterexample

does have utterance meaning, but it would be tendentious to maintain that an utterance

that cannot be understood should possess utterance meaning.

Now, certainly a counterfactual could be constructed in which my utterance does

have utterance meaning; something like, "if I had told my interpreters that when I say 'I

fleuvenate' I intend to be interpreted as meaning 'I agree', then my utterance would have

had utterance meaning". However, a counterfactual like this will always involve a case

in which my justified belief that I will be interpreted as I intend is true, not false. The

---

[21]There are of course other ways to develop a justified belief which is in fact false. Perhaps I have been the victim of a prank.

existence of such a counterfactual is not a response to the counterexample. The counterexample works because the JP2 says only that the speaker's belief about how they will be interpreted must be justified, not necessarily true, which allows for the construction of a case which qualifies as having utterance meaning according to the JP2, but is clearly gibberish. Pointing out that if the speaker's justified belief were true then their utterance would have utterance meaning does not affect the strength of the counterexample.

There is, of course, the temptation to simply add to the JP the requirement that the speaker's belief that his utterance would be interpreted as they intended must be justified *and* true. Adding a truth requirement would avoid the second counterexample, but it would reintroduce the first. The first counterexample provides a case where the speaker had a justified belief that they would be interpreted as they intended and their utterance did have utterance meaning, though their justified belief is not true. Adding a truth requirement to the JP would once again force us to deny that the utterance in the first counterexample has utterance meaning.

Clearly, something has gone wrong with Davidson's proposal, but what exactly is it? The problem is in part exaggerated by Davidson's failure to address it properly. The footnote is a sign of the recognition that the AP is insufficient, but it does not provide a solid course to any changes that might rescue it. If Davidson recognized that the AP was problematic, he ought to have said more about what might be done to adjust it. It may also be that the AP is not salvageable, but it would be premature to make that claim.

The fundamental problem with the AP and the two versions of the JP appears to be that the success or failure of a speaker's intention does not determine the meaning of our utterances. It would be far too strong to say that this has nothing to do with meaning;

20

it seems clear that intentions are closely tied to the meaning of our words. Determining the degree to which intention affects utterance meaning would take much more work which cannot be pursued here. However, the first counterexample shows that intention and intention satisfaction is not sufficient for utterance meaning.

Davidson's fundamental mistake was attempting to amend the notion of a satisfied intention by including justified beliefs and modal terms instead of entirely reconsidering the relation of satisfied intention to utterance meaning. The notion of a satisfied intention may appear to explain how some, perhaps many, of our utterances possess meaning, but it falls apart when forced to consider counterexamples in which intentions fail but utterance meaning remains. If appeals to the satisfaction of intentions cannot explain how our utterances, as a whole, possess meaning, then it has failed to serve its purpose.

What can be said about utterance meaning at this point? At least that the actual interpretations of a speaker's utterance, as well as the speaker's beliefs about how their utterances would be interpreted under certain conditions, no matter how justified or true those beliefs may be, are not sufficient for the presence of utterance meaning. So far the claims made here have been negative. If Davidson's proposals are in need of adjustment, in what ways should they be adjusted? What can be said about a positive account of utterance meaning?

At the beginning of this essay Davidson's rejection of subjectivist and conventionalist accounts of utterance meaning was outlined. Seeing that Davidson's account has run out of steam, it is worth reexamining Davidson's arguments against subjectivist and conventionalist accounts, since they claim to provide the sort of account we seek. If Davidson's attempt to explain utterance meaning in the absence of

subjectivity or conventions has failed, perhaps these previously rejected accounts hold the

solution to our problem. If so, Davidson has made two mistakes: rejecting either

conventionalist or subjectivist accounts, and claiming that the satisfaction of intentions

can provide a standard that the previous two theories seemingly did not. In the next

chapter, we will re-examine the arguments for a subjectivist theory of utterance meaning

and assess the strength of Davidson's arguments against them.

Chapter Two

Since we have at least for now abandoned Davidson's attempts to provide a standard of utterance meaning in terms of a satisfied intention, we will now examine Davidson's project of rejecting the subjectivist proposal, specifically Davidson's proposed alternative to subjectivism. According to the subjectivist, utterance meaning can be accounted for entirely in terms of speaker meaning. What a speaker's words mean is just the meaning that the speaker has in mind. If this position seems untenable given the preceding chapter, notice that when communication succeeds, it is typically because our words express the meaning we have in mind. It is not unreasonable to suppose that communicative success is not a matter of assigning the correct meanings to other people's words as Davidson claims, but a matter of determining what other people intend to communicate independently of the words they happen to use.

Davidson's rejection of the subjectivist account in part results from his rejection of the traditional account of subjective mental content. In a sense, the subjectivist account is not terribly different from Davidson's own proposal. Davidson's proposal involves intentions, which are also subjective states. The difference is in the content of the intention; Davidson's intentions are about how one's *words* should be interpreted, while subjectivist intentions are about a *content* that is to be communicated and not about any particular words that are capable of expressing that content. The meaning of a speaker's words, as far as the subjectivist is concerned, has only to do with what is "in their head", and Davidson fundamentally disagrees about what it means for something to be "in one's head."   One way to see the basic difference between Davidson's and subjectivist positions is to look at how each proposal views communicative failure. For Davidson, communicative failure can be the result of a speaker's failure to say something

meaningful. As we saw in the first chapter, a speaker can fail in this regard by speaking in a way that cannot be interpreted. If there is no meaning that can be assigned to a speaker's utterance, that utterance has no meaning and communication fails. In these cases, the speaker has "failed to give the meaning to her words that she intended the interpreter to catch", regardless of what content she may have had in mind or what intentions she possesses. Communication failed because the *speaker* failed to say anything meaningful.

The matter of speaker error is simpler for the subjectivist. As long as a speaker has some content in mind which he intends to communicate with an utterance, that utterance has meaning. Since a speaker's words mean simply what the speaker has in mind, it is impossible that a speaker could have a certain content in mind but fail to use the right words. Whatever words the speaker uses can convey that content solely by virtue of the fact that it is the content that the speaker has in mind. If I say, "blass morpul tront?" with the content "Would you take the dog for a walk?" in mind, then that is the content expressed by my utterance; that is what my utterance of "blass morpul tront" means.[22] If my interpreter fails to understand my utterance, then it is not because my utterance was without meaning, but because my interpreter failed to grasp that meaning. Davidson's proposal entails that "someone can't mean something by his words that can't be correctly deciphered by another" something that the subjectivists must consider possible if not frequent.[23]

---

[22]There is a qualification that must be made here. Davidson does not deny that an utterance of something like "blass morpul tront" *could* be a meaningful utterance. What Davidson denies is that having some content in mind is *sufficient* for making a meaningful utterance.

[23] Donald Davidson," Knowing One's Own Mind," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 28.

The difference in views of communicative failure highlights the difference in views of communicative success. For Davidson, speakers grant meaning to their utterances only when both speakers *and* interpreters perform their roles correctly. For the subjectivist, speakers grant meaning to their utterances simply in virtue of their own behavior and mental content, and then it is up to the interpreter to correctly identify that meaning.[24] So utterance meaning for Davidson is determined by an interaction between speaker and interpreter while utterance meaning for the subjectivist is determined solely by the speaker, potentially in the absence of any interpreter. For Davidson, meaning and communicative success are connected; without the existence of some communicative success, our utterances are not meaningful, since the meaning of our utterances is tied up with the presence of interpreters. For the subjectivist, the meaning of our utterances is *completely independent of any success that others have in interpreting them*; the subjectivist's conception of utterance meaning would exist even if there had never been a single instance of successful communication.

One significant result of this difference between Davidson and the subjectivists is that the existence of a "private language" is a possibility for the subjectivist. That is, it is possible that there could be a language that is only interpretable by the speaker of that language, a possibility that is excluded from the proposals made by Davidson as well as conventionalists. According to Claudine Verheggen, Davidson claims that "a solitary person, that is, a person who has been socially isolated from birth…could not have a

---

[24]Of course, it is not always possible for an interpreter to correctly identify the meaning that a speaker has in mind when making an utterance. The point is just that for the subjectivist, it is possible to make an utterance that cannot be interpreted by another.

language."[25] If it were shown that a private language was not possible, it would certainly

be strong evidence against the subjectivist proposal.

Davidson is not alone in denying the possibility of a private language. For

reasons that are similar to Davidson's, conventionalists like Kripke also affirm that

language is necessarily social. The concern is one initially developed by Wittgenstein:

language requires the possibility of error. Both Davidson and the conventionalists

believe that social interaction is necessary in order for error, and thereby language, to be

possible. For his part, Davidson sees a language as a set of truth conditions that specify

the truth conditions for every possible sentence in the language.[26] To speak a given

language is to *correctly* follow the truth conditions for that language, and this entails that

there must be a capacity for error; without the possibility of error there can be no

possibility of correctness. This is one purpose served by the process Davidson calls

"triangulation" :, to explain how social interaction creates the possibility of error. These

claims will not be evaluated here; they simply illustrate the motivation for Davidson's

proposals.

The argument against a private language is not, however, Davidson's primary

reason for rejecting the subjectivist proposal. The fundamental basis of Davidson's

rejection is epistemological. Subjectivism requires that mental contents can be described

without any reference to objects and events external to the mind, and Davidson insists

that mental content cannot be described without such references. While it is of course

significant that Davidson rejects a basic premise of subjectivism, it is in a way

superfluous to the discussion here. The aim here is not to defend subjectivism but to

---

[25] Claudine Verheggen, "Triangulating With Davidson," *Philosophical Quarterly* 57:226 (2007), 96-103.
[26]Of course this will be an infinite set due to the recursive nature of language.

critique Davidson's project, and we can still provide such a critique without contesting Davidson's rejection of subjectivism. Once Davidson has rejected the subjectivist account of the relation between mental content and the meaning of our utterances, he must provide an alternative proposal, and this alternative will be the subject of the remainder of this chapter.

As we have seen, Davidson's account of meaning presented in the first chapter presupposes a social setting; there must be both a speaker and an interpreter in order for there to be meaningful utterances. Davidson is aware, though, that even once subjectivism is rejected it cannot simply be assumed that language requires more than one person as his theory presupposes, and he makes an attempt to demonstrate that a private language cannot exist. Why then, does Davidson consider a private language to be an impossibility? Because it is only in the presence of other people that a speaker can be said to be talking *about* anything. In order to understand this, it must be understood that for Davidson, part of what determines the meaning of a speaker's words is the way in which the speaker is typically *caused* to utter those words.

> "What a person's words mean depends in the most basic cases on the kinds of objects and events that have caused the person to hold the words to be applicable: similarly for what the person's thoughts are about."[27]

Let's relate this claim to the previous chapter. The job of an interpreter is to understand how the speaker intends her utterance to be understood. To figure out how a speaker intends their utterances to be understood, we must figure out how their words are connected to the world. As Davidson puts it, "to designate a language as one being

---

[27] Donald Davidson, "Knowing One's Own Mind" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001) , 37.

spoken requires that utterances be matched up with objects and events in the world"[28].

This is true not only for interpreters, but speakers as well. In order to utter something that is *interpretable*, the utterance must match something in the world that caused that utterance.[29] The interpreter's job is that of discovered which language is being spoken, and that can be achieved by matching words with objects and events in the world, but obviously those connections between speaker and world must exist in order to be speaking a language which can then be interpreted. That which does not exist cannot be discovered.

Davidson's argument against the possibility of a private language depends on this picture of the relation between language and world described above. Of course, the sort of connection between language and world that has just been described is not sufficient for the existence of language; it is just a necessary condition for the existence of language. Neither is it the only necessary condition of the existence of language. We will forgo any disputes over the necessity of a causal connection between language and world in order to contest a different necessary condition of language that Davidson introduces. In addition to being causally connected to the world, Davidson requires that there be some specifiable cause to which a speaker's utterance is a response. Since we have granted that part of what it is to be interpretable is to provide one's interpreters with utterances that can be "matched up with objects and events in the world", it seems that indeed there must be some specifiable object or event which can be matched to a speaker's utterance in order for that utterance to be interpretable and thereby meaningful.

---

[28] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 120.
[29] The phrase, "something in the world" is being used in a very broad sense here, encompassing more than just material objects.

Specifying the cause of a speaker's utterance seems simple since we do it so regularly, but there is a problem involved in describing a creature's responses to the world. We naturally class other people's (and other animals') behavior as directed towards certain parts of the environment: "Rob swerved his car to avoid hitting that fence," or "My dog isn't eating his food because he doesn't like chicken." It seems unquestionable that fences, food, cars, and chicken are all objects in the world to which Rob and my dog are reacting. (Davidson's examples are of a child calling a table a 'table' and a dog salivating in response to a bell). The question is:

> "Why say the stimulus [of the dog's salivation] is the ringing of the bell? Why couldn't it be the vibration of the air close to the ears of the dog – or even the stimulation of its nerve endings?...Why not say the same about the child: that its responses are not to tables but to patterns of stimulation at its surfaces, since those patters of stimulation always produce the response, while tables produce it only under favorable conditions?"[30]

This is essentially the difficulty, introduced by Devitt and Sterelny, referred to as the "qua problem".[31] There is a (sometimes quite long) causal chain between myself and anything in the world to which I can be said to react. These chains involve many (relatively) small events between light waves, particles of air, and the molecules that make up the objects I am reacting to. When describing my reaction, why describe it as a reaction to one part of the causal chain than another? Why say that when Rob swerves, he swerves in response to the fence, and not in response to the patterns of stimulation on his eyes? His behavior would not change if the causal chain led to some fence-illusion creating device rather than an actual fence. It is the patterns of stimulation that are

---

[30] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 118.
[31] Sterelny Devitt , *Language and Reality: An Introduction to the Philosophy of Language* (Oxford: MIT Press, 1999), 90.

necessary for Rob to behave in the way he does, not the fence. Why then, do we say that Rob swerved because of the fence instead of saying he swerved because of the patterns of stimulation on his eyes? There is no clear answer. Given this incapacity to individuate between Rob's behavior and the world, how can we give an answer that will accommodate the description, "Rob swerved to avoid the fence"? Davidson suggests that a process he calls "triangulation" provides the answer.

Triangulation is a familiar process. In geographical contexts, it determines the distance of a third on the basis of the distance between two other points and the angles formed by creating a triangle that joins all three points. With enough information about the properties of first two points, we can discover properties of a third point. In Davidson's version of triangulation, the first two points are creatures and the third point to be determined is the cause of the creatures' reactions. Davidson claims that when it comes to a solitary creature, one that has never interacted with others like it, there can be no answer to the question, "to what is it responding?" or "what is causing its behavior?" due to the difficulties of the qua problem. If this is true, then since for Davidson the ability to interpret another's speech depends on the ability to individuate the causes of the speaker's utterances (at least in the most basic cases), a solitary creature cannot be interpreted, and therefore cannot speak a language. When a creature triangulates with another like it, however, Davidson suggests that the cause of the creature's reaction can be individuated in a way that allows for the possibility of language. The qua problem threatens Davidson's thesis that the ability to individuate the cause of a speaker's (verbal) reaction is a necessary condition of utterance meaning, but triangulation promises an answer.

Instead of a single dog salivating when a bell is rung, imagine that there are two dogs that salivate when the bell is rung. In this situation, the causal chains that are responsible for each dog's behavior have something in common: the ringing of the bell. Davidson says, "We may think of it as a form of triangulation: each of the two people is reacting differently to sensory stimuli streaming in from a certain direction. Projecting the incoming lines outward, the common cause is at their intersection."[32] Both dogs have the same reaction: salivation. Since they react in the same way, it is reasonable to suppose that they are reacting to the same thing, and the only commonality between the causal lines is the bell itself. The intersection of causal lines is the basis of triangulation, and since Davidson considers triangulation a necessary condition of language it makes sense that he would claim that, "Communication begins where causes converge."[33] Convergence of causal lines can pick out the ringing of the bell as 'the' cause of the dog's behavior rather than the patterns of stimulation on the dog's sense organs.

Davidson submits to the qua problem in the case of a solitary creature. He concedes that if a single dog were salivating in the presence of a ringing bell, without any previous social interaction with other dogs (or similar animals), then there would be no individuating any part of the total cause as 'the' cause of the dog's reaction. If a particular cause cannot be individuated in these simple cases, then there is little hope of individuating 'the' cause of a solitary speaker's utterances and therefore little hope of interpreting them.

The qua problem poses a particular difficulty for Davidson. In his words,

---

[32] Donald Davidson, "Three Varieties of Knowledge," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 213.

[33] Donald Davidson, "A Coherence Theory of truth and Knowledge," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 151.

"Without other people with whom to share responses to a mutual environment, there is no answer to the question what it is in the world to which we are responding."[34] This is somewhat imprecise; it is not that there is *nothing* to identify as a cause of the solitary creature's reaction. There is the total cause, the entire causal chain that extends back in time perhaps infinitely, that is responsible for the creature's reaction. The problem is that the total cause will not suffice as a candidate for 'the' object or event to which the creature is reacting. There are simply too many competing objects and events contained within the total cause for it to be of any use when two or more creatures try to "each correlate their own reactions to external phenomena with the reactions of the other.[35] To what should I correlate the reactions of the other? The objects or events that are currently present, or ones that are long past? The most distant cause, or the closest one? Davidson continues, "The reason has to do with the ambiguity of the concept of cause. It is essential to resolve these ambiguities, since it is, in the simplest cases, what causes a belief that gives it its content."[36] In order to know what a speaker's utterances mean, I must know what they intend, what they believe, what they desire, and so on. In short, I must know the content of their mental states. The problem is that in order to have mental content there must be some part of the world that is 'the' (relevant) cause of that content. The qua problem suggests that the cause of mental content cannot be individuated, and without mental content, there can be no interpretation or meaning as far as Davidson is concerned.

There are two ambiguities that stand in the way of individuating the causes of mental content. "The first ambiguity concerns how much of the total cause of a belief

---

[34] Donald Davidson, "The Emergence of Thought" *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 129.
[35]Ibid.
[36]Ibid.

32

is relevant to content…The second problem has to do with the ambiguity of the relevant stimulus, whether it is proximal (at the skin, say) or distal."[37] Resolving the second ambiguity is a matter of individuating a particular object or event from the total cause; for example, individuating a table from the many other objects and events that are contained in the total cause like the making of the table or the light patterns that reflect off the table, etc. Once the object or event has been individuated, there is a further difficulty. Which parts or aspects of the object or event are relevant and which are simply incidental? Is it the color that is causing the reaction? The shape? The function? Davidson claims that triangulation resolves both ambiguities.

Though triangulation is in the end a linguistic project, the most basic form of triangulation (which we will call "pure" triangulation) can involve non-linguistic creatures who are not, and never will be, capable of language or propositional thought. According to Davidson, even pure triangulation resolves the two ambiguities mentioned above, which must be resolved if we are to proceed to the more complex cases of triangulation that involve linguistic creatures capable of propositional thought. Pure triangulation, then, should be able to provide us with some narrower "part or aspect of the total cause" which can serve as 'the' cause of the creatures' reactions. [38]   The convergence of causal lines, as Davidson sees it, does just that. The convergence of causal lines gives us a subset of the total cause that is narrow enough to serve as 'the' cause of the creature's reaction; it resolves the second ambiguity.

Peter Pagin brings up a simple problem that begins to complicate the picture. Pagin puts it like this: "If an event c is a common cause of two other events, a and b,

---

[37]Ibid., 129-130.
[38] Ibid.

then any further event *d* which is a cause of *c* is also a common cause of *a* and *b*."[39] If

the reactions of creature **A** and creature **B** are caused by the appearance of a butterfly,

there is more than a single common point between the long causal chains that resulted in

the reactions of creatures **A** and **B**. They also share the long causal chain of events,

extending back perhaps infinitely, that led to the appearance of the butterfly and is

equally common to the causes of both creatures' reactions. The causal chains that led to

each creature's reactions then share not only the appearance of the butterfly, but the entire

causal chain that led to the appearance of the butterfly as well. The picture is more like

that represented by figure 1 than figure 2.
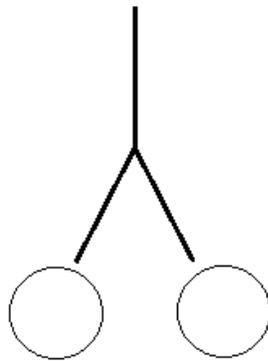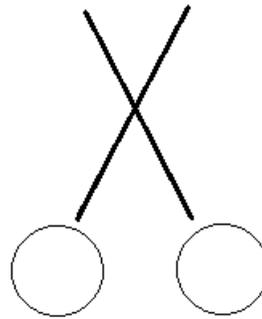

Figure 1   *Figure 2*

The concern, then, is that as Pagin says, "there is no such thing as *the* common

cause."[40] Instead of being able to pinpoint 'the' cause as the common cause, we are left

with a further question: "which one of the many common causes is 'the' common cause?"

Pagin is willing to forgive Davidson on this matter, saying it is reasonable to assume that

the *closest* common cause is "most plausibly what the talk of *the* intersection of causal

lines comes down to."[41] I am not convinced that this problem is so forgivable.

---

[39] Peter Pagin, "Semantic Triangulation" in *Interpreting Davidson*, (Stanford, Calif.: CSLI Publications, 2001), 202.
[40] Ibid.
[41] Ibid.

If we are going to take the qua problem seriously, and it is clear that Davidson thinks we should, then it is troubling that even once the convergence of causal lines provides a narrower space there is still so much room for ambiguity. Triangulation was supposed to individuate an area in which '*the*' cause of the creature's reaction is present by isolating some part or aspect of the total cause. Triangulation was required because the entirety of the total cause was too big; there was no part or aspect that could reasonably serve as 'the' stimulus. But even now that we have identified a causal intersection between the triangulating creatures a significantly large space remains, a space that contains a great many different objects and events. Claudine Verheggen summarizes the basic problem nicely, "To endow one's utterance with meaning, one is supposed somehow to connect it to its typical cause. But, Davidson asks, which cause is that?"[42] Isolating the common causes from the total cause does not get us appreciably closer to answering the question "which cause is that?"

Suppose that we allow Davidson the rather ad-hoc qualification that it should be the *closest* common cause which serves to individuate 'the' cause of a creature's reaction. There are then further complications that the following example should illustrate. Imagine a world with only two creatures, Creature A and Creature B. Creature A and Creature B have never interacted before, but happen to be sitting relatively close to one another when a breeze comes up and knocks a coconut loose from a tree next to Creature A, causing Creature A to let out a yelp. At the same time a completely distinct cause, say a minor tremor, has set loose a coconut from a tree next to Creature B, causing Creature B to let out a similar yelp. Each creature hears the other's yelp and reacts to that as well,

---

[42]Claudine Verheggen, "Triangulating With Davidson," *Philosophical Quarterly* 57:226 (2007), 98.

turning to look for the cause of the noise. Here is a case of two interacting creatures that

have been caused to have similar reactions. Each reaction has a total cause, a causal

history that extends back indefinitely, and those causal histories intersect. The causal

lines of both creatures' reactions have a cause in common, but it is not the breeze or the

tremor, since the breeze was unique to Creature A's causal chain and the tremor was

unique to Creature B's causal chain. It happens that the closest common cause of both

reactions is an animal that deposited the seeds that grew into the coconut trees. There is a

common cause, but it is in the distant past, before either creature existed.

Clearly in a case like this, individuating the closest *common* cause does not help

to individuate a stimulus that is relevant to the present reactions of both creatures.

Something more is needed to individuate 'the' cause of the creatures' reaction in this

case. Neither should we say that they are not triangulating. They interact with each

other, each responding to one another's yelps that were in turn caused by two different

falling objects, and the causal histories of those reactions share a common object or

event. The problem is that the common cause is so distant that it is of no use.

Now imagine another scenario with Creatures A and B. Both creatures have

access to a machine that presents different images to each of them. To either creature, the

machine sometimes presents images of a threat, and sometimes images of their favorite

food. Both creatures react similarly to similar stimuli; they react to the image of the

threat by crouching and to the image of the food by salivating. Both creatures have

access to the reactions of the other creature as well as the machine's images, but the

machine is inconsistent. Sometimes it displays the same image to both creatures;

sometimes it displays different images to each creature and neither creature has access to

the images of the other. There is a closest common cause to the reactions of both

creatures in reasonable temporal proximity to both creatures: the machine. But again, the individuation of the machine as the closest common cause does not seem to help individuate the cause that is relevant to both creatures' reactions.

Davidson has summarized this "primitive" triangulation by saying "It involves two (or as always, more than two) creatures reacting to the same scene, event, or object, and correlating the other's reaction with the observed external stimulus."[43] In the case described here, the creatures would have little to correlate, since the responses of the other are not consistent with the images that appear to the creature. The individuation of the closest common cause (the machine) might help us to identify 'the' cause of the creatures' responses, but then we must wonder whether that would do any good in this case, since neither creature will correlate the behavior of the other with the machine. Perhaps this is not a true instance of triangulation, but Davidson does little to illustrate why that would be. If there is more to triangulation than the existence of a common cause, which these examples suggest there must be, Davidson appears not to acknowledge it.

So not only does the isolation of the set of common causes accomplish little without relying on the questionable assumption that the closest common cause is the relevant cause, even once that assumption is granted there are cases in which the closest common cause appears not to be the relevant cause. But suppose these problems are resolved as well. Could Davidson then proceed? If we are somehow convinced that talk of the 'closest common cause' is unproblematic, how does it solve the ambiguity of the qua problem? Well, one problem was that the total cause of a creature's reaction was too

---

[43] Donald Davidson "Comments on Karlovy Vary Papers" in *Interpreting Davidson* (Stanford, Calif.: CSLI Publications, 2001).

wide to give any meaning to the idea that there was some particular object or event to which the creature was reacting. Talk of the closest common cause was intended to narrow the field, so to speak, by picking out a particular part of the total cause as 'the' cause. Suppose we grant Davidson the (albeit problematic) notion of a closest common cause. The question I wish to now raise is, "why should 'the' cause be common?"

The resolution of the qua problem demands some means of individuating a particular cause from the total cause that can serve as 'the' cause of a speaker's mental content. As Davidson puts it, determining "how much of the total cause of a belief is relevant to content."[44] Pure triangulation accomplishes this by making the closest common cause 'the' relevant cause. It must be the *closest* common cause because there are many common causes; the notion of a common cause is still too broad to pick out 'the' cause. Davidson appears to have either ignored the fact that there are many common causes or believed that the 'closest' qualification is not worth making explicit, because he says nothing that acknowledges any further ambiguity among 'common' causes. After describing the triangular relationship he moves directly to talk of 'the' common cause.

Neither would it be useful to avail ourselves of the fact that there is often only one common cause that is perceivable by both triangulating creatures. If there is a causal intersection between two humans at, say, a television, clearly the only cause perceived by both people is the television and its images, not the light waves that pass from the television screen to each viewer. It is tempting to then say that we are not individuating the 'closest' common cause but the common cause that is present to both creatures.

---

[44] Donald Davidson, "The Emergence of Thought" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 129-130.

Helping ourselves to "objects that are present to a creature" introduces a bit of question-begging, however, for this assumes that the object that is "present" to the creature can be individuated, and wasn't the original problem of this very sort? If we could individuate 'the' object *present* to a creature there would be little difficulty individuating 'the' cause of the creature's reaction. If we take the qua problem seriously, we must resolve the ambiguities it presents before our talk of 'the' object present to a creature is warranted. So we should stick with the notion of common causes, and if so we need the 'closest' qualification to disambiguate 'the' cause from the many common causes.

It would seem, though, that it is the "closest" qualification that is doing all the work for Davidson. When first introducing the qua problem, Davidson writes in "The Second Person," "if we must choose, it seems that the proximal cause of the behavior has the best claim to be called the stimulus, since the more distant an event is causally from its perceiver, the more chance there is that the causal chain will be broken."[45] Davidson makes this claim with the intention of eventually denying it, but it is this sort of reasoning that would have to come into play to substantiate talk of the "*closest* common cause." If it is justifiable to individuate 'the' cause from among the many *common* causes on the basis of proximity, then it should be justifiable to use proximity alone to individuate 'the' cause from the total cause. We could then skip the business of triangulation and common causes and simply say that 'the' cause of a creature's reaction is simply the *closest* cause, rather than the closest *common* cause.

The common cause was supposed to pick out a section of the total cause that would serve as 'the' cause of the creature's reaction. Then we observed that in order to

[45] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 118.

pick out a small enough section, we must use the closest common cause. Now I suggest that the closest cause serves to individuate a sufficiently small section of the total cause just as well as the closest common cause. The closest cause will individuate a single object or event that can serve as 'the' cause of a creature's reaction that is relevant to its mental content, and this is all that it needed to solve the qua problem.

The qua problem required an answer because it was an obstacle to Davidson's notion of mental content. If there is no particular cause which can be individuated, then there is no particular cause that can serve as 'the' cause of a creature's reaction and therefore there can be no mental content. But now we have an answer to the qua problem: 'the' cause of a creature's reaction is the cause that is causally closest to the reaction. While it may solve the qua problem, it simply passes on the difficulty to other aspects of Davidson's theory.

Triangulation serves several purposes for Davidson. Aside from the qua problem, there is an additional problem regarding the possibility of error. One fundamental difference between Davidson and the subjectivists, as mentioned at the beginning of the chapter, is that for Davidson the capacity to make an error is a necessary condition of meaning. As far as Davidson is concerned, in order to mean something there must be a difference between saying something meaningful and simply intending to say something meaningful; we must be able to make mistakes. Triangulation is supposed to help explain the capacity for error as well as solve the qua problem.

Triangulation explains the capacity for error, claims Davidson, by showing how creatures can come to possess the concept of objectivity. For Davidson, the concept of objectivity, the idea that some things are the way they are independent of our minds, is necessary if a creature is to have any mental content at all. It is not practical to contest

this claim here, so we will just grant it to Davidson. Verheggen summarizes the position, saying that, "one cannot have a language without having beliefs, and that one cannot have beliefs without having the concept of belief."[46] Furthermore, one cannot have the concept of belief without having a concept of error. Therefore, one cannot have a language without having a concept of error. "Having a belief demands in addition appreciating the contrast between true belief and false, between appearance and reality, mere seeming and being."[47] Davidson summarizes this connection between belief and error with the concept of objectivity, which he contends is made possible by triangulation.

Triangulation, Davidson says, allows for the possibility of the concept of objectivity "by providing (as I have cautiously put it) a 'space' in which awareness of the possibility of error can take root."[48] Creatures can become aware of the possibility of error by correlating the behavior of a second creature with certain objects or events in a common space, as creatures do in triangulation. Expectations are created, and eventually the creature comes to expect the behavior and object or event to come together they expect the presence of the object or event when they observe the correlated behavior of the second creature and they expect the behavior of the creature when they observe the presence of the object. The chance to recognize the possibility of error arises when those expectations are not met; when the creature behaves as if the object is present but it is not, or when the object is present but the creature does not behave as expected.

---

[46] Claudine Verheggen, "Triangulating With Davidson," *Philosophical Quarterly* 57:226 (2007), 100
[47] Donald Davidson, "Three Varieties of Knowledge," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 209.
[48] Donald Davidson, "The Karlovy Vary Papers," in *Interpreting Davidson* (Stanford, Calif.: CSLI Publications, 2001), 293.

The sort of situation Davidson has in mind is this: "A learned reaction can be observed in certain monkeys which make three distinguishable sounds depending on whether they see a snake, an eagle, or a lion approaching; the other monkeys, perhaps without seeing the threat themselves, react to the warning sounds in ways appropriate to the different dangers, by climbing trees, running, or hiding."[49] This is a case of pure triangulation in which expectations are created, and then can fail to be met; a monkey might make the sound that signals the approach of a lion when there is no lion approaching, or a snake might approach but no sound is made. This, Davidson claims, is the sort of situation that makes objectivity possible.[50]

Objectivity was not a concern when discussing pure triangulation since the creatures involved had no mental content or language for which objectivity would be required. Pure triangulation was only designed to "give meaning to the idea that the stimulus [of a creature's reaction] has an objective location in a common space"[51] But the previous observations suggest that pure triangulation has failed to give meaning to that idea. Davidson believes that the triangulation necessary to solve the qua problem provides the 'common space' that is necessary for objectivity, but we have seen that the solution to the qua problem does not put 'the' cause of a creature's reaction in a common space but in the causally closest space to the creature's reaction. In fact, the space that is causally closest to a creature's reaction is almost surely not common. So Davidson has another impasse. The qua problem can be a resolved, but not in the way that Davidson

---

[49] Donald Davidson, "The Emergence of Thought" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 128.

[50] Davidson does not claim that the monkeys have the concept of objectivity, since they do not possess propositional thought or language. He only claims that this is the sort of triangulation that makes the concept of objectivity possible in creatures that *are* capable of language and propositional thought.

[51] Donald Davidson, "The Second Person," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 119.

wanted; namely, one that provides for the possibility of objectivity. If the solution to the qua problem does not also explain how objectivity is possible, Davidson needs another way to explain how creatures might come to have the concept of objectivity that he claims is a necessary condition of propositional thought and language. It seems as though the triangulation that Davidson has in mind still fits this bill, however. The story involving the creation of expectations and the failure to have those expectations met could still suffice as an explanation of how objectivity is possible. But could an explanation of that sort proceed without a second creature? I think that it can.

Expectations can exist even when considering a solitary creature because triangulation requires just three *points*, only one of which need be a creature. Imagine a world with only a single creature. The creature instinctually explores its environment, seeks out food, shelter, and the other necessities of life. After a while, the creature begins to notice that heavy rain and rising water levels often occur around the same time. The correlation is so strong that the creature begins to develop expectations around these two events; when there is heavy rain the creature expects the water level to rise, and when the water level rises the creature expects heavy rain. These expectations are met   throughout most of the creature's life, but eventually they fail to be met. Eventually there is a heavy rain but the water level does not rise, or the water level rises but no rain follows. Here we have a triangle composed of causal lines; one line goes between creature and water level, a second between creature and rain, and the third between rain and water level, as we see in figure 3.
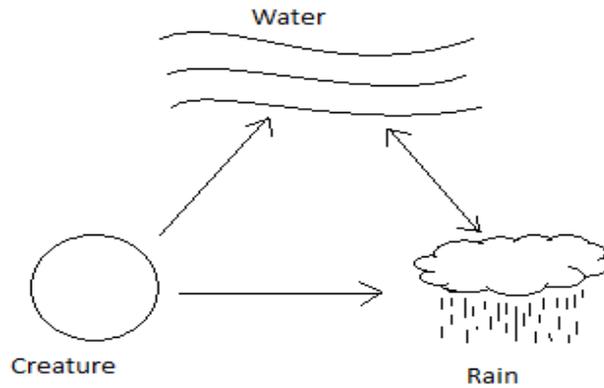
Figure 3

This sort of situation would seem to provide the same sort of "space in which the awareness of the possibility of error can take root" as Davidson's situations that involve two creatures rather than one. [52] It is not necessary that the creature realizes that a mistake has been made or even have the capacity for such a realization, since Davidson's purpose is not to create a case in which the concept of objectivity actually arises but to point to a situation that must occur if the concept of objectivity is ever to arise.

So there have been two obstacles to Davidson's project thus far: the qua problem and the problem of objectivity. The qua problem threatened Davidson's explanation of mental content by obscuring the relationship between mental content and the cause of that content. The problem of objectivity arose when the resolution of the qua problem was found to be unsuitable for explaining the possibility of objectivity. Both problems now have solutions but they are not the solutions proposed by Davidson. So what is the consequence of introducing these alternative solutions?

If these alternative solutions are accepted, then they no longer provide for the possibility of language as Davidson conceives of it. Davidson's notion of language and

---

[52] Ibid.

thought is one of a necessarily social phenomenon that requires a common, external

cause of each speaker's utterances. The alternative solution to the qua problem does not

provide us with a common external object that can be individuated as 'the' cause of a

creature's reaction and the alternative solution to the problem of objectivity does not

require more than one creature. If Davidson wishes to support his concept of language,

he will need different arguments than the ones posed here.

Subjectivism has hardly been vindicated, but neither has Davidson. Davidson in

parts motivates his own proposal by highlighting the insufficiencies of subjectivist and

conventional positions. If by rejecting subjectivism Davidson has simply changed the

problems rather than resolved them, then he has presented us with little motivation to

abandon subjectivism in favor of his proposal. However, Davidson may still acquire

some motivation for his position from a solid rejection of conventionalism, which we will

examine in chapter three.

Chapter Three

In the first chapter we examined Davidson's own account of utterance meaning and found it to be lacking. Davidson's account was motivated by his rejection of subjectivist and conventionalist proposals, so in the second chapter we examined Davidson's attempts to deal with the problems that resulted from his rejection of subjectivism. The third and final chapter will be a similar examination of Davidson's attempts to reject conventions as the basis for a standard of utterance meaning.

We can start our examination with the similarities between Davidson and the conventionalist. Both positions deny the possibility of a private language; that is, they both claim that language is possible only in a social environment, though for different reasons. The conventionalist affirms the social nature of language because conventions require at least two people. Davidson affirms the social nature of language because he believes the necessarily social process of triangulation is required for the existence of mental content and objectivity, as was noted in the previous chapter.

Both Davidson and the conventionalists believe that utterance meaning is dependent on this social environment, but to different degrees. For the conventionalist, utterance meaning is entirely determined by the purely social function of convention; the meaning of a speaker's words depends on the linguistic conventions of the other people in the speaker's community. A formalization of the conventionalist standard of utterance meaning was presented in chapter one:

> **S**'s utterance **U** means that **P** if and only if there exists a convention that utterances of **U** mean that **P.**

Given this definition, it is clear that a speaker could intend to mean something by an utterance but still fail to give their words the meaning they intend. So in contrast to the subjectivist, the conventionalist agrees with Davidson that there must be the possibility of error for there to be meaning. So words, for the conventionalist, as for Davidson, cannot simply mean whatever a speaker intends to them mean by his utterance; there are conditions external to the speaker that play a part in determining the meaning of an s utterance. The conventionalist recognizes the importance of external conditions but in Davidson's eyes places too much importance on those external factors. What words mean is determined more by the verbal habits of the speaker and her fellow speakers rather than by the speaker's intentions. But if Davidson quarreled with the subjectivists for failing to recognize the importance of external factors to utterance meaning, his dispute with the conventionalists is the result of a failure to recognize the importance of *internal* factors to utterance meaning. As we saw in the first chapter, intentions are central to Davidson's theory of utterance meaning, and a theory that fails to give appropriate weight to those intentions is surely insufficient in his eyes.

One of the main tasks when arguing for or against the conventionalist position is defining convention itself. David Lewis' essay "Languages and Language" provides the definition that Davidson attacks and this paper will employ.[53] In his essays "Communication and Convention" and "A Nice Derangement of Epitaphs."[54]Davidson evaluates the claim "that the meaning of a word is conventional, that is, that it is a convention that we assign the meaning we do to individual words and sentences when

---

[53]David Lewis, 'Languages and Language', in *Language, Mind, and Knowledge,* (University of Minnesota Press, Minneapolis, 1975).
[54] Davidson, "A Nice Derangement of Epitaphs" in *The Essential Davidson*, (New York: Oxford University Press, 2006).

they are uttered or written."[55]]He goes on to recount Lewis' definition of convention,

which has six conditions:

> "a convention is a *regularity R* in action, or action and belief, a regularity in which more than one person must be involved. The regularity has these properties: (1) Everyone involved conforms to R and (2) believes that others also conform. (3) The belief that others conform to R gives all involved a good reason to conform to R. (4) All concerned prefer that there should be conformity to R. (5) R is not the only possible regularity meeting the last two conditions. (6) Finally, everyone involved knows (1)-(5) and knows that everyone else knows (1)-(5), etc."

These conditions will serve as the definition of convention that Davidson has in mind

when he claims that conventions are not a necessary condition of utterance meaning.

Lewis' definition fits nicely with the conventionalist standard of utterance meaning stated

above, as well as our intuitions about conventions. When I use the word "table", my

utterance means table only if conditions 1-6 are met, where R is "meaning table by the

utterance 'table.'"

This is the target of Davidson's critique. To oversimplify, the idea is that the word

"table" means table if that is what the word is regularly used to mean. The idea seems to

reflect the way language works. We could mean anything by the words we use, "table"

could mean chair and "chair" could mean balloon. What makes table the meaning of

"table" is that we agree, as a community of English-speakers, that it does. This

agreement is not necessarily a conscious, explicit agreement among speakers of a

language but a code of behavior to which everyone who speaks the language conforms.

The position is appealing, but Davidson's position on the role of convention in

language is subtle. He does not deny that convention can, and often does, play a role in

our interpretation of others' utterances. "Knowledge of the conventions of language is

thus a practical crutch to interpretation, a crutch we cannot in practice afford to do

---

[55] Davidson, "Communication and Convention". 13

without – but a crutch which…we can in the end throw away." His claim is only that conventions are not *necessary* for understanding the speech of others. Davidson's argument is often a simple modus tollens:

> P1: If conventions are necessary in order to determine the meaning of our utterances, then utterances that do not accord with established conventions cannot be meaningful.
>
> P2: There are meaningful utterances that do not accord with any established convention.
>
> C: Therefore, conventions are not necessary to determine the meaning of our utterances.

I must say that I find the argument convincing, but it is important to recognize the narrow nature of Davidson's claim. Davidson's argument eliminates only a particular role of conventions, which leaves room for convention to play a necessary role in communication. Since it is easy to interpret Davidson's claim as a much broader claim about *any* role that convention might play, it is necessary to take some time to outline the limitations of Davidson's argument. In "A Nice Derangement of Epitaphs", Davidson focuses on the existence of malapropisms, which for Davidson are essentially instances of speech that diverge from conventional speech, whether intentionally or by mistake. For example, "the pinochle of success", "hitting the nail right on the thumb", and "chickens always come home to roast."[56] For Davidson, "What is interesting is the fact that in all these cases the hearer has no trouble understanding the speaker in the way the speaker intends."[57]

---

[56] Donald Davidson, "A Nice Derangement of Epitaphs" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press), 252.

[57] Ibid.

As Davidson notes, it's easy to see how an understanding of the speaker is reached: "The absurdity or inappropriateness of what the speaker would have meant had his words been taken in the 'standard' way alerts the hearer to trickery or error; the similarity in sound tips him off to the right interpretation."[58] Davidson points to malapropisms as instances of successful communication that fail to accord with any rule or convention.

There is no way, then, for a malapropism to satisfy the first condition of convention: that everyone conforms to that use of the word. In fact, malapropisms can fail to be a regularity at all since they are sometimes unintentional and immediately corrected. If a speaker accidently uses the word "roast" rather than the word "roost", it is quite possible that the speaker may realize their mistake and fail to ever make that same mistake again. In such a case, the malapropism could be an entirely unique utterance, a usage of a word that has never occurred before and will never occur again. If such an utterance is understood and communication is successful, then it would seem to be a case in which communication succeeds but convention is absent.

Davidson also evidences premise 2 with theoretical principles. Davidson contends that Interpretation, , is an inevitable part of communication. We can observe this by considering how an individual might function if they had somehow learned a language in isolation from a dictionary and a book of formation rules for constructing sentences. Let us call him Dictionary Man. Dictionary Man appears to have a grasp of the language and his utterances are all easily understood, but it would not take many interactions with typical speakers of the language before Dictionary Man encountered a word that he had

---

[58] Ibid.

not learned from the dictionary, at which point he would be forced to interpret the speaker in order to understand the speaker's utterance.

Interpretation appears to be an unavoidable practice among language speakers, no matter how similar their speech habits may be. Davidson points to names. New names appear at a considerable rate, and the introduction of a new name requires many instances of interpretation; every time a speaker is faced with an unfamiliar name interpretation must take place. But interpretation is not something which can be explained by appeals to convention. The process of interpretation, claims Davidson, is like any other method we use for navigating the world; as new evidence comes in, we adjust our theories to accommodate the new data. Interpretation operates in a similar way.

Competent speakers and interpreters come to a particular communicative exchange with certain expectations (Davidson calls them "theories") of how the other will speak, just as scientists come to particular experiments with certain theories of how their experiment will proceed. We develop those expectations on the basis of many things; clothing, geographical location, gender, education, etc. [59] Often, however, those expectations or theories are not completely, if at all, accurate. I may come to a linguistic exchange with the expectation that I will be presented with some dialect of English, but instead hear Spanish. When our theories are found not to be accurate, we adjust them to accommodate the new data. When a speaker uses a language that I did not expect him to speak, I do not try to interpret him using the language I thought he would speak, I adjust my interpretation. When a speaker makes an utterance that does not conform to our

---

[59] Theories in this sense are merely descriptive. Davidson is not making claims about what speakers and interpreters know or how they think. Talk of 'theory' is simply a way to describe what the speakers and interpreters do.

expectations (as malapropisms tend to do), "the interpreter adjusts his theory so that it yields the speaker's intended interpretation."[60]

Davidson's theoretical point is that interpretation itself depends on our ability to construct and adjust these sorts of theories, and that ability cannot be characterized by rules or conventions. "There are no rules for arriving at [a correct interpretation], no rules in any strict sense, as opposed to rough maxims and methodological generalities."[61] There is no rule or regularity to be observed in previous scientific experiments that can tell us how to adjust our theories in the face of new data that does not conform to those previously observed regularities. Analogously, there is no rule or regularity to be observed in previous communicative exchanges that can determine the meaning of new utterances that do not conform to our previously observed regularities. Interpretation, like theory construction, is not the kind of practice that is regular, pre-established, arbitrary, or to which everyone conforms. In short, it is not conventional.

Davidson does not provide anything that he considers to be definitive proof of this thesis, but the ball is now in the conventionalist court. I believe that the claim which Davidson's argument is designed to reject, that the meaning of our words is determined by convention, is indefensible. For that reason, I defend a slightly different claimthat conventions must exist in order for communication to proceed. If Davidson's claim that interpretation is just another form of theorizing is correct, then I would need to show that the theorizing that takes place during interpretation also involves conventions, and Davidson is highly skeptical of such a demonstration. "There is no more chance of regularizing, or teaching, [the process of interpretation] than there is of regularizing or

---

[60] Davidson, "A Nice Derangement of Epitaphs" in in *Subjective, Intersubjective, Objective* (New York: Oxford University Press), 258.
[61] Ibid., 265.

teaching the process of creating new theories to cope with new data in any field – for that is what this process involves."[62]

This ability to interpret the speech of others, not according to any established, rule-based practice but by "wit, luck, and wisdom"[63], makes communication possible in the absence of convention and therefore in principle unnecessary for communication. Conventional practices must be shared, according to Davidson, but if interpretation is possible in the absence of convention then communication can proceed without any shared practices that can be considered conventional. Davidson imagines a world with two speakers, one that speaks Sherpa and one that speaks English. There could be no conventions between the two, since neither speaker follows the same rules. Yet as long as both people have the (non-conventional) ability to interpret the speech of the other there is no reason that communication could not proceed as it does between two English speakers or two Sherpa speakers.

This is perhaps not the most persuasive example since we often assume that being able to interpret a language means that you can speak that language. Imagine instead that the two people are physically different such that they are not capable of making any of the same sounds. There is then no way for either person to speak the language of the other since it is physically impossible, and again there can be no conventions between them. But still, as long as each person has the ability to interpret the speech of another, communication could proceed. It is not a condition of interpretation that one be able to mimic the other. By the same token, two people might use different names to refer to the same object, yet each could understand the names used by the other. The conventionalist

---

[62] Ibid.
[63] Ibid.

53

would have to show that in order to interpret you, I must share your patterns of speech; I must speak like you.

I believe that P1 of Davidson's argument, as I interpret it, is true and that the argument as a whole is valid. If conventions are a necessary condition of meaning, then meaningful utterances that truly do not rely on conventions could not exist, and if P2 is true as well then the conclusion follows necessarily. I merely wish to show that while *shared* conventions are not sufficient for communication, conventions still play a necessary role in the process of communication. Davidson's main evidence is the existence of malapropisms, utterances that fail to agree with existing conventions but are nevertheless understood. But there is a difference between utterances that are meaningful even when no linguistic conventions exist, and meaningful utterances that simply do not *agree* with existing linguistic conventions.

If Davidson intends to refer to utterances that are meaningful in the absence of *any* linguistic convention, then he would surely be right that the existence of such utterances would provide a counterexample to the conventionalist claim. However, it is not clear that such utterances exist. Davidson is quite willing to admit that linguistic conventions exist and function centrally in everyday communication. "I want to urge that linguistic communication does not require, though *it very often makes use of, rule-governed repetition.*"[64] Conventions, Davidson says, are "a practical crutch to interpretation, a crutch we cannot in practice afford to do without."[65] So Davidson is not claiming that we could communicate as we do now without establishing some, if not many, linguistic conventions. He is simply saying that linguistic conventions are not

---

[64] Davidson, "Communication and Convention," *Synthese* 59 (1984) 3-17. Italics are mine.
[65] Ibid.

sufficient for communication, and that malapropisms are evidence that communication can proceed even when the meaning of our words is not dependent on convention.

In what sense do the malapropisms presented by Davidson fail to agree with linguistic conventions? Well, they are certainly uninterpretable by convention alone, since there are no conventions that suffice to interpret utterances like "the pinochle of success". But the existence of this sort of utterance only shows that knowledge of conventions alone is insufficient for interpretation, not that it is unnecessary. If knowledge of conventions is *unnecessary* for the interpretation of malapropisms then not only will malapropisms be uninterpretable through conventions alone, the interpretation of malapropisms will not require any knowledge of linguistic conventions whatsoever.

Conveniently, Davidson provides a description of the process involved in interpreting a malapropism:

> "The hearer realizes that the 'standard' interpretation cannot be the intended interpretation; through ignorance, inadvertence, or design the speaker has used a word similar in sound to the word that would have 'correctly' expressed his meaning. The absurdity or inappropriateness of what the speaker would have meant had his words been taken in the 'standard' way alerts the hearer to trickery or error; the similarity in sound tips him off to the right interpretation."[66]

So when convention alone does not provide a coherent interpretation, the interpreter begins to search through similar-sounding 'standard' utterances that would make sense in the given context. But this only seems possible given that most, if not the rest of the utterance is 'standard', or bears some relation to a 'standard' utterance. The title of Davidson's essay, "A Nice Derangement of Epitaphs", is interpretable because of the close similarity in sound to an utterance that does follow linguistic conventions, "a nice arrangement of epithets". It seems not to matter whether the malapropism is "a nice

---

[66] Ibid., 252

derangement", "a mice merangement", a "lice de-fangment", or any other similar-sounding utterance. When the utterance no longer bears that similarity to a conventional utterance, or bears a closer similarity to a different conventional utterance, the ability to interpret 'correctly' wanes. It is difficult for me to imagine ordinary cases in which utterances like, "snarf fingle farble", or "a wife's estrangement" should be interpreted as "a nice arrangement".[67]

Despite the explanation of the interpretive process that appears above, Davidson seems to think that the similarity in sound is irrelevant. "Similarity of sound is not essential to the malaprop. Nor for that matter does the general case require that the speaker use a real word; most of 'The Jabberwock' is intelligible on first hearing."[68] Frankly, this is a claim which I have little hope of understanding, especially since this appears to be the totality of Davidson's thoughts on the matter.

I think Davidson is right to say that though similarity in sound is one of many ways to come to a 'correct' interpretation, and that "there are many other ways the hearer might catch on."[69] The question is whether any of these ways do not in *some* way depend on knowledge of conventions, and Davidson fails to provide any evidence that interpretation which does not involve something conventional actually occurs. Analogously, there is an established convention among my roommates and I that an empty roll on the counter means that we are out of toilet paper, but the fact that we can spot *and* correct mistakes merely demonstrates that we are very good at recognizing patterns, not that convention needn't play any role in that process.

---

[67] Whether or not we can imagine a fantastic scenario in which this sort of thing occurs is irrelevant at this point, since the current concern is with Davidson's empirical evidence. Consideration of his theoretical evidence will follow.
[68] Ibid., 252
[69] Ibid., pg. 251

In addition to empirical claims, and perhaps more centrally, Davidson provides a theoretical justification for P2. As with the empirical evidence, my concern is to show that while Davidson's theoretical evidence may demonstrate that conventions are not sufficient for interpretation, it fails to show that conventions are unnecessary. In order to convincingly demonstrate that conventions are unnecessary, Davidson needs a case in which interpretation occurs but no conventions are present.

Davidson's finds his strongest theoretical evidence in what I will refer to as the "limit case". It is a case put forward by Davidson that is designed to show that interpretation and communication can proceed in the absence of shared convention:

> "If you and I were the only speakers in the world...we could understand one another, though each of us followed different 'rules' (regularities). What would matter, of course, is that we should each provide the other with something understandable as a language. This...does not involve following shared rules or conventions. It might even be that because of differences in our vocal chords we couldn't make the same sounds, and therefore couldn't speak the same language."[70]

Davidson's evidence for the possibility of such a case is somewhat sparse. In "The Second Person", he simply states that "I know of no argument that shows that under such circumstances communication could not take place."[71] In "A Nice Derangement of Epitaphs", Davidson states even more baldly that "in principle communication does not demand that any two people speak the same language."[72] In "Communication and Convention" Davidson comes the closest to providing real evidence for this claim, saying "Different speakers have different stocks of proper names, different vocabularies, and

---

[70]Davidson, "The Second Person," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press), 114
[71] Ibid.
[72] Davidson, "A Nice Derangement of Epitaphs," in *Subjective, Intersubjective, Objective* (New York: Oxford University Press), 257.

attach somewhat different meanings to words."[73] While surely true, it remains unclear

whether these differences are simply instances of non-conventional speech or instances of

non-conventional speech imbedded in an otherwise conventional-dependent language.

I believe that it is at least questionable whether or not such a case is possible, but I

do not wish to debate it here. Instead, I will grant that communication is possible in the

limit case and will simply suggest that convention plays a role even in the limit case. I do

not claim that it plays the sort of role that it is the primary target of Davidson's critique,

the claim that for communication to take place there must be a convention that "speaker

and hearer mean the same thing by uttering the same sentences"[74]. I only intend to

defend the claim that convention cannot, even in principle, be done away with entirely if

communication is to take place. Therefore the limit case is the relevant case to examine,

for if no convention exists and communication proceeds, my thesis would clearly be

false. However, if convention can be shown to play a necessary role even in the

communication that takes place in the limit case, then Davidson's insistence that

convention is unnecessary for language itself will become that much more suspect.

What reason would we have for believing that conventions could exist at all in the

limit case? Davidson is quite confident that the existence of conventions in the limit case

is not possible, claiming that "convention requires conformity on the part of at least two

people", and "The analysis clearly requires that there be at least two people involved,

since convention depends on a mutually understood practice."[75] If this interpretation was

secured then perhaps Davidson's claims would follow without question, but this is not an

---

[73] Davidson, "Communication and Convention," *Synthese* 59 (1984) 3-17
[74] Davidson, "Communication and Convention" *Synthese* 59 (1984) 3-17.
[75] Ibid.

interpretation shared neither by myself nor by David Lewis. Lewis in fact explicitly

claims that convention can exist in a solitary individual:

> "The isolated man conforms to a certain regularity at many different times. He knows at each of these times that he has conformed to that regularity in the past, and he has an interest in uniformity over time, so he continues to conform to that regularity instead of to any various alternative regularities that would have done about as well if he had started out using them. He knows at all times that this is so, knows that he knows at all times that this is so, and so on. We might think of the situation as one in which a convention prevails in the population of different time-slices of the same man."[76]

Lewis' suggestion that the concept of convention may be applied to the behavior

of even a single individual has significant implications for our discussion, since it would

provide for the possibility of conventional behavior in the limit case. Davidson, though

he clearly resists the conclusion, might be forced at least part of the way towards

accepting Lewis' claim. In his treatment of the limit case, Davidson allows that both

speakers have linguistic regularities to which they conform, but focuses on the fact that

the two speakers do not follow the *same* regularities. Since we have granted that speaker

and interpreter need not share linguistic conventions for communication and are instead

concerned with whether conventions must play any role at all in communication, the

observation that the speakers in the limit case do not *share* any linguistic regularities is

somewhat irrelevant to the present discussion.

The more interesting observation is that both speakers *have* linguistic regularities,

regardless of whether they are shared or not, and Davidson has, though perhaps

grudgingly, made it clear that if either speaker is to be interpreted, they must speak with

some regularity. Davidson says, "carrying out this intention [to provide the other with

something understandable as a language], while it may require a degree of what the other

---

[76] David Lewis, "Languages and Language" pg. 148

perceives as consistency, does not involve following shared rules or conventions."[77] Again, communication may not require following *shared* rules or conventions, but it does not follow that communication does not require following rules or convention.

It is worth pausing on Davidson's ambiguous claim that speaking a language requires providing the other with "a degree of what the other perceives as consistency." citation here It seems to me that the mere appearance of consistency or regularity is insufficient for interpretation. For interpretation to succeed, one's utterances must maintain a significant degree of regularity, since the regularity perceived by the interpreter must correspond to an actual regularity in the habits of the speaker.

So Davidson was somewhat inaccurate when he said that a speaker must present an interpreter with just "a degree of what the other perceives as consistency." To be precise, a speaker must present an interpreter with speech patterns that are actually consistent with one another. In order to communicate, the speaker must continue to mean the same thing by the same words. At least, that is, until interpretation is successful. Once communication succeeds, the speaker may modify the meanings of their words and interpretation begins anew, but a speaker must use words in the same way until interpretation succeeds *if* communication is to occur. To do otherwise is to sabotage the process of interpretation itself.

If a speaker must continue to mean the same thing by the same words if interpretation, and thereby communication is to occur, then it seems that something much like convention is in place even if that regularity is not shared with another person. Davidson claims that "speaking a language cannot depend on speaking as someone else

---

[77] Donald Davidson, "The Second Person" in *Subjective, Intersubjective, Objective* (New York: Oxford University Press, 2001), 114.

does (or as many others do)."[78] True, speaking a language does not depend on speaking as others do, but it does depend in part on the previous speech of the speaker. The speaker must continue to speak as they have before. Until, of course, interpretation succeeds. Davidson himself agrees that a regularity of this sort must be present, so even Davidson's limited requirements for language are consistent with Lewis' claim that convention can persist in a solitary individual.

The consistency that Davidson finds necessary for interpretation provides the room for convention to exist in the limit case. If either speaker is to have any hope of being interpreted, they must provide the other with something interpretable as a language, which in turn must contain certain regularities. Those regularities, across time, satisfy the definition of convention that Lewis lays out and Davidson accepts: a *regularity R* in action that has these properties: (1) The speaker (at different times) conforms to R and (2) believes that he previously conformed to R. (3) The belief that he previously conformed to R gives him a good reason to conform to R now. (4) He previously and currently prefers that he should conform to R. (5) R is not the only possible regularity meeting the last two conditions. (6) Finally, the speaker knows (1)-(5) and knows that he knew (1)-(5) previously, etc.

All these conditions are met by a single speaker in the limit case. Condition (1) is met by virtue of the fact that the speaker conforms to a regularity over time. Condition (2) is met because the speaker believes that he is conforming to this regularity, a belief he must have if he intends to "provide the other with something understandable as a language.[79] Satisfaction of condition (3) is conditional upon the preference for successful

---

[78] Ibid., 115.
[79] Ibid., 114

communication, but if that desire is present then he has good reason to conform to his previous regularity since that conformity is necessary for communication. Satisfaction of condition (4) is equally conditional on the desire that communication succeed, but as long as that desire is present his preference will be for conformity. Conditions (5) and (6) are satisfied in the same way in the solitary case as they are in the case where conventions are shared between people.

Lewis' suggestion that we consider the convention as "shared" between different time slices of the same person seems to fit the agreed upon definition without any significant modifications. The only changes to the definition have been made to better reflect that the regularity is shared across time rather than between different people. I see no reason to deny that conventions are present in the limit case as well.

Lewis writes, "Conventions are regularities in action, or in action and belief, which are arbitrary but perpetuate themselves because they serve some sort of common interest. Past conformity breeds future conformity because it gives one a reason to go on conforming."[80] All these conditions are met in the limit case; the linguistic behavior of the people in the limit case is regular, arbitrary, the regularity is perpetuated because of an interest that the speaker has over time, and their past behavior gives them a reason to continue behaving in the same way.

Davidson then does not have evidence that communication can proceed in the absence of convention, though he may have evidence that communication can proceed in the absence of conventions that are shared between speaker and interpreter. Granting Davidson that conventions need not be shared between speaker and interpreter entails that convention cannot play a particular sort of role in the process of communication;

---

[80]Lewis, David. "Languages and Language". (From pg. 135 in language and mind)

62

conventions cannot be something that speaker and interpreter share that makes communication possible. This is, of course, not the only role that convention can play in the process of communication. The weakest version of my claim would be that convention is a necessary but not sufficient condition of successful communication, and this is what Davidson fails to deny.

The consequence is that Davidson's arguments have at best shown that the conventionalist ought to amend their proposal to account for the broader role that convention plays in the process of communication. Evidence that a proposal should be modified, however, is not evidence that a proposal should be rejected.

So where does Davidson's proposal for a standard of utterance meaning now stand? In a way, we have been working backwards. We first examined Davidson's proposal for a standard of utterance meaning that emerged from his rejection of subjectivist and conventionalist proposals. In the end, we were unable to find a way to adjust the proposal that could account for the existence of several counterexamples. Given the flaws of Davidson's own account, we turned to the other two dominant positions, subjectivism and conventionalism, to see if the demand for a new account could be substantiated. Subjectivism, while suffering from its own difficulties, was not easily disposed of. Davidson's attempts to construct alternatives to the subjectivist position failed to amount to a problem-free account that warranted a move away from subjectivism. The attack on conventions ran into similar problems; though Davidson managed to carve out a particular part of language that is not conventional (namely interpretation), his argument is still far from showing that the role of conventions is not a necessary and important one that requires some attention in a sufficient account of utterance meaning.

Given these results, it would be unwise to grant Davidson's account any sort of prominence in relation to the other existing proposals. Not only does his own account fail in important respects, it is unclear that the alternative proposals do not provide an equally viable strategy. As is often the case, there is likely some truth to each position and it is our duty to isolate those truths so that they may be retained in future versions or entirely new proposals. I believe Davidson is right to attempt a position that gives significant weight both to the importance of mental content and the necessity of social interaction, and I think that any competing theory must address both components.

As to the prospects of Davidson's theory overcoming the objections raised here, I see no reason to think that they are irreconcilable. In fact, the shortcomings of his theory may turn out to be only superficial, but they are shortcomings nonetheless and require correction, superficial though they may be. If anything, this paper aspires to provide a clarification of the difficulties involved in such a project and to organize them in a way that promotes further advances.

In a more positive sense, I think the limitations of Davidson's proposal illustrate the importance of simply having the intention to mean something by an utterance. The counterexamples posed in the first chapter point to cases in which the existence of such an intention, whether it is satisfied or not, appear to be sufficient for utterance meaning. That is not to say that having such an intention is always sufficient for utterance meaning, but that it is at the very least a central component of utterance meaning. What the other components are is a matter of further discussion, but it seems clear that our intuitive understanding of what it takes to utter meaningful words is to a large extent dependent on the intention to be understood in a certain way.