

The Philosophy of Language

Lecture Eight

Davidson's Theory of Meaning

Rob Trueman
rob.trueman@york.ac.uk

University of York

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

Grice's Project

- Grice wanted to reduce meaning to intention
- By uttering u , speaker S means that p
 - (i) By uttering u , speaker S intends to induce the belief that p in their audience A
 - (ii) S intends A to recognise intention (i)
 - (iii) S intends A to form the belief that p at least in part because A recognises intention (i)
- x sentence-means that p iff x is *regularly* or *conventionally* uttered with the speaker-meaning that p

Presupposing Intention

- Grice's reduction presupposes that the notion of intention is in good order
- This is something that someone like Quine would surely doubt
 - A Quinean would likely add *intention* to their list of inter-definable but ungrounded concepts, along with *analyticity*, *synonymy*...
- There is also an epistemological question
 - How can we recognise that someone has the complex intentions that Grice appeals to without already presupposing that we can understand the sentences that they use?
- Some Griceans have tried to get around this by distinguishing between **analytic** priority and **epistemological** priority
- However, it isn't entirely clear how good this response is

Problems for the Analysis of Speaker-Meaning

- Grice's analysis of speaker-meaning faced two kinds of counter-example
- **The analysis is too weak**
 - There are some cases which meet all of the conditions in Grice's analysis, but which are not cases of speaker-meaning
- **The analysis is too strong**
 - There are some cases of speaker-meaning which do not meet all of the conditions in Grice's analysis

A Problem for Grice's Analysis of Sentence-Meaning

- Grice's analysis of sentence-meaning has trouble accounting for the meanings of sentences which have so far gone unuttered
- The trouble is that Grice's analysis of sentence-meaning does not seem to be compositional
 - To say that meaning is compositional is to say that the meaning of a sentence is determined by the meanings of the expressions it is built out of, and the way it is built out of them
- Grice can try to introduce compositionality by thinking of the intentions involved in speaker-meaning as being built out of smaller sub-intentions
- But it is not obvious that Grice can force compositionality into this account in this way

This Week: Davidson's Theory of Meaning

- This week, we will look at a Davidson's theory of meaning
- Davidson put the compositionality of meaning front and centre, and didn't try to squeeze it into his theory as an afterthought
- Davidson was also deeply concerned about the epistemological problem surrounding intentions and beliefs
- Although Davidson thought that recognising someone's intentions and beliefs was importantly **related** to understanding what they were saying, he did **not** think that you could just take intentions for granted in the way that Grice did

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

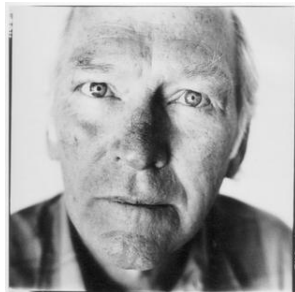
Radical Interpretation

The Principle of Charity

An Objection

Conclusion

- Davidson was a very influential philosopher in the 20th Century
- He wrote landmark papers in the philosophy of action, mind and language
- Quine was Davidson's mentor
- Davidson does disagree with Quine on a number of points, but it is easy to see Quine's influence on his work



Donald Davidson

What is a Theory of Meaning?

- Davidson wants to figure out how to make a **theory of meaning**
- What would a theory of meaning for a language L do?
 - (i) It would *somehow* present the meaning of each sentence in L
 - (ii) It would also make clear *how* the meaning of a sentence in L depends on the meaning of its parts
- Our theory of meaning needs to do (ii) to make it clear that English is **learnable**

Learnability

- As we mentioned last week, there are **infinitely many** English sentences
- Most of those sentences have not yet been uttered, but they all have a meaning in English
- What is more, if I uttered one of these previously unuttered sentences, you would be able to figure out what it means
- This amazing ability of yours can be accounted for if we make our theory of meaning compositional
 - The meaning of a sentence is determined by the meanings of its parts
 - You can understand a new sentence you have never come across before because you know what the words mean, and you see how they've been put together

What A Theory of Meaning *Doesn't* Need To Do

- A theory of meaning *does not* have to **reduce** meaning to something more basic
- A theory of meaning for English just needs to tell you what each English sentence means, and how the meaning of a whole English sentence depends on the meanings of its parts
- We might hope that such a theory will help us to understand how *meaning* relates to various other concepts, like *belief* and *intention*
- But it does not need to **define** *meaning* in terms of these concepts!

What A Theory of Meaning *Doesn't* Need To Do

- A theory of meaning for English *does not* have to be a theory which every speaker of English explicitly knows
- We might instead say that the theory of meaning is something that every English speaker **implicitly** knows
 - At some sub-conscious level, the brain is operating in accordance with the theory
- Or we might give up on the idea that every speaker knows the theory at all, whether explicitly or implicitly
 - The theory of meaning is just an abstract theory which presents the meaning of every English sentence, and explains how the meaning of a whole sentence depends on the meanings of its parts

Meanings as Entities?

- I said that a theory of meaning for language L must *somehow* present the meaning of each sentence in L . But how exactly?
- In other words, what should the **outputs** of this theory be?
- You might think that the outputs should have this form:
 - x means m
 - 'Socrates is wise' means *the fact that Socrates is wise*
- On this view, meanings are a type of entity, or object, which you can refer to with a singular term, like '*the fact that Socrates is wise*'

No, Not Meanings as Entities!

- Davidson does not like this whole approach
- He gives detailed arguments at the start of his paper 'Truth and Meaning' and also in 'True to the Facts', but the jist of it is just this:
 - We have no idea what these weird entities called 'meanings' are meant to be, or how they are meant to work
 - Trying to build a theory of meaning which treats meanings as entities is therefore bound to be a waste of time
- As a result, he insists that we should not expect a theory of meaning to give us outputs like this:
 - x means m

From Meaning to Truth

- Instead, Davidson thinks that we should explain what a sentence means by giving its truth-conditions:
 - 'Socrates is wise' is true in English iff Socrates is wise
 - 'La neige est blanche' is true in French iff snow is white
- Sentences **say that** things are a certain way, and when you know the truth-condition of a sentence, you know how that sentence says things are

The Quest for a Theory of Truth

- According to Davidson, the quest for a theory of *meaning* is really the quest for a theory of *truth*
- This was a very exciting conclusion for Davidson, because in the 1930s, Alfred Tarski had developed a detailed, formal theory of truth
- What is more, Tarski's theory was compositional
 - The truth-condition of a whole sentence depends on the semantic properties of its parts
- Davidson's hope was that he could re-purpose Tarski's theory of truth as a theory of meaning
- But to understand what that might mean, we need to take a quick look at Tarski's work

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

Theories of Truth

- Philosophers have presented lots of different things that have been called 'theories of truth', but we have a particular type of theory in mind
- For our purposes, a theory of truth for a language L is simply a theory which somehow presents the **truth-conditions** for each sentence in L
- Such a theory is not required to answer any deep, metaphysical questions about what truth is

Tarski's Achievement

- Tarski was the first person to figure out how to develop a theory of truth for even moderately complex languages
- Unfortunately, it is well beyond our means to give a full account of Tarski's achievement here
- Instead, I will just try to give a taste of what a theory of truth might look like



Alfred Tarski

Simplish — A Simple Language

- **Names**

a, b

- **Predicates**

F, G

- **Connectives**

\neg, \wedge

- **Example Sentences**

Fa, Ga, Fb, Gb, $\neg Fa, Fa \wedge Gb, \neg(Fa \wedge Gb), \dots$

A Theory of Truth for Simplish

- How would we make a theory which attributed truth-conditions to all of the sentences in Simplish?
- Although Simplish is very simple, it is already complex enough to generate infinitely many sentences
- We cannot, then, go through them and give them truth-conditions one by one
- Instead, we will have to come up with a series of compositional rules which will then fix a truth-condition for each and every sentence

A Theory of Truth for Simplish

- 'a' refers to Barry Lee
- 'b' refers to Mary Leng

- An object satisfies 'F' iff it has a moustache
- An object satisfies 'G' iff it is a philosopher of mathematics

- The result of writing a predicate followed by a name is true iff the referent of that name satisfies that predicate

- If S is a sentence, then $\neg S$ is true iff S is false

- If S_1 and S_2 are both sentences, then $S_1 \wedge S_2$ is true iff S_1 is true and S_2 is true

A Theory of Truth for Simplish

- ' Fa ' is true iff Barry Lee has a moustache
- ' Gb ' is true iff Mary Leng is a philosopher of mathematics
- ' $\neg Ga$ ' is true iff it is not the case that Barry Lee is a philosopher of mathematics
- ' $Fa \wedge Gb$ ' is true iff Barry Lee has a moustache and Mary Leng is a philosopher of mathematics

How to Choose between Theories of Truth

- I've just presented you with **one** theory of truth for Simplish, but there are many other possibilities
 - For example, we might take 'a' to refer to Tom Stoneham rather than Barry Lee
- How do we pick between them?
- Tarski introduced a general test for picking between different theories of truth

A Test of our Theory

- Imagine that Simplish was a real language, spoken by real people
- We could then ask how to translate the sentences of Simplish into English
- Suppose we found out that '*Fa*' should be translated as 'Barry Lee has a moustache'
- In that case, Tarski insists that the correct theory of truth for Simplish should entail the following:
 - '*Fa*' is true iff Barry Lee has a moustache
- As it happens, our theory of truth does imply that, so it would pass this test

Convention T

- More generally, Tarski said that a decent theory of truth for a language L must pass **Convention T**
- For any sentence s in L , the theory of truth must imply a result of the following form:
 - s is true in L iff p
- Importantly, ' p ' must be the **meta-linguistic translation** of s
 - When we are theorising about a language, we call the language we are theorising about the **object-language**, and the language we are using to do the theorising the **meta-language**
 - So in our earlier example, Simplish was the object-language, and English was the meta-language

Testing a Theory of Truth for French

- Suppose we were trying to construct a theory of truth for French, using English as our meta-language
- We would need to make sure that we got results like this:
 - ‘La neige est blanche’ is true in French iff snow is white
 - ‘L’herbe est verte’ is true in French iff grass is green
 - ‘Socrate est sage’ is true in French iff Socrates is wise
- Any theory of truth for French (formulated in English) which did not give these results would fail Convention T, and so would not be a good theory of truth

Testing a Theory of Truth for English

- Convention T becomes even simpler when the meta-language happens to be the same as the object-language
- Suppose we were trying to construct a theory of truth for English, using English as our meta-language
- We would need to make sure that we got results like this:
 - 'Snow is white' is true in English iff snow is white
 - 'Grass is green' is true in English iff grass is green
 - 'Socrates is wise' is true in English iff Socrates is wise
- Obviously, these look absolutely trivial, but that is just because the object-language and the meta-language are one and the same

Summing Up

- As I said, we haven't even begun to do justice to Tarski's achievement here
- All of his hard work was focussed on figuring out how to deal with quantifiers, and we haven't taken a moment to look at them
- These are the two things to bear in mind about Tarski:
 - (1) He figured out how to construct **compositional theories of truth**, which use the semantic properties of sub-sentential expressions to fix the truth-conditions of whole sentences
 - (2) He introduced Convention T as a way of testing whether a proposed theory of truth was a good one

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

Davidson's Programme

- Davidson wanted to co-opt a Tarski-style theory of **truth** for a given language as a theory of **meaning** for that language
- Tarski himself was very sceptical of the idea that we could actually build one of his theories of truth for real, **natural** languages
 - Tarski focussed on well behaved **formal** languages, and thought that natural languages were just too messy to build theories of truth for
- It is still an open, controversial question whether we can give a Tarskian theory of truth for, e.g., English
- Davidson did lots of work to try to show that it was possible, but we will not look at any of that work today

Why use a Theory of Truth as a Theory of Meaning?

- **Compositionality**

- Davidson emphasised that a theory of meaning must be compositional if we are to account for our ability to understand novel, previously unuttered sentences
- Tarskian theories of truth are compositional, and so are perfectly suited to this job

- **Convention T**

- A theory of meaning must somehow present the meaning of each sentence
- A good Tarskian theory of truth will satisfy Convention T, and so will use translations of each object-language sentence to express their truth-conditions

An Obvious Problem

- Tarski used Convention T to test his theories of truth, which presupposes the concept of translation
 - For each sentence s , a good theory of truth for L will imply ' s is true in L iff p ', where ' p ' is the meta-linguistic **translation** of s
- But if we are now trying to use a Tarskian theory of truth as our theory of meaning, we surely cannot just presuppose the concept of translation!
 - To say that one sentence *translates* another is to say that they **mean** the same thing
 - So a theory of meaning shouldn't just assume the concept of translation; it should help us understand that concept!

We cannot Rely on Convention T as a Test!

- Davidson was well aware of this problem, and insisted that we couldn't use Convention T to test a theory of meaning
 - Importantly, Davidson was not **rejecting** Convention T
 - He still thought that a good theory of meaning should satisfy that Convention
 - He was just saying that we couldn't use Convention T to **test** whether a theory of meaning was any good
- We need to come up with a whole new way of testing our theories of meaning!

Radical Interpretation

- Imagine you are a linguist who discovers a long lost civilisation in a remote part of the world
- You decides to try to interpret their language, LongLostish, into English
- Unfortunately, the language is totally unrelated to any language that you know and so you have to work totally from scratch
- This is known as **radical interpretation**
 - Davidson discusses this in his paper 'Radical Interpretation', which is this week's seminar reading

Interpretation versus Translation

- Davidson's radical **interpretation** is obviously a lot like Quine's radical **translation**, but there is difference
- Quine is interested in translating one language into another, which is a matter of pairing each sentence in the first language with a sentence in the second
 - 'La neige est blanche' translates as 'Snow is white'
- Davidson is interested in interpreting a language, which is a matter of somehow explaining what the sentences in that language mean
 - 'La neige est blanche' says that snow is white

Gava gai

- One day, you notice a rabbit run past, and one of the people from the lost civilisation shouts, 'Gava gai!'
- You notice this happens often: *whenever* someone sees a rabbit, they shout 'Gava gai!'
- You note the following conjecture in about LongLostish
 - 'Gava gai' is true in LongLostish iff a rabbit is nearby
- You test this conjecture by saying 'Gava gai' a few times when a rabbit is in plain sight, and people seem to agree that what you said is true

Some More Evidence

- Then some time later, you notice a monkey run past, and one of the people from the lost civilisation shouts, 'Bava gai!'
- You see this happen a few more times, and conjecture:
 - 'Bava gai' is true in LongLostish iff a monkey is nearby
- You also notice that when a rare type of white monkey is nearby, people say 'Zagga bava gai', and so you conjecture:
 - 'Zagga bava gai' is true in LongLostish iff a white monkey is nearby
- You test these conjectures by trying the sentences out yourself in the right circumstances, and people seem to agree with what you say

Developing a Theory

- So far, you haven't really begun developing a theory of truth for LongLostish
 - You have made some conjectures about what the outputs of that theory should include, but you do not yet have a theory which gives you those outputs
- However, you now know enough to start trying to construct a theory:
 - An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is white
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- This theory will give you all the truth-conditions you have already checked

Testing the Theory

- An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is white
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- This little theory implies:
 - 'Zagga gava gai' is true in LongLostish iff a white rabbit is nearby
 - So you could test this theory further by saying 'Zagga gava gai' when you see a white rabbit, and seeing if the locals seem to agree with what you say

Semantic Holism

- According to Davidson, this is the general pattern of how we cook up a theory of truth, and then test it
- You start by focussing on how speakers use certain sentences, and try to assign them truth-conditions which would fit the way they are used
- After you have done this for enough sentences, you can start noticing patterns between the sentences, and use these patterns to make conjectures about what the parts of the sentences mean

Semantic Holism

- These conjectures come together to make a theory, which will offer new predictions about the truth-conditions of sentences you haven't yet checked
- If those predictions fit the way that the speakers use those sentences, then your theory is confirmed and you can carry on developing it
- If those predictions do not fit the way that speakers use those sentences, then your theory is disconfirmed, and you might need to go back and revise it

Semantic Holism

- This leads Davidson to endorse a position known as **semantic holism**
- According to semantic holism, it does not really make sense to ask what a given sentence means in isolation
- We can only ask what a single sentence means in the context of a whole language
- That's because the only way of investigating whether we have interpreted a given sentence correctly is by using that interpretation to make predictions about what **other** sentences mean, and then checking whether those predictions are right
 - There is clearly an affinity here between Davidson's **semantic holism** and Quine's **confirmational holism**

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

Another Obvious Problem

- I just said that we can check whether our theory of truth for a given language is a good theory by checking whether it fits with the way that its speakers use the sentences in that language
- But what does it mean to say that our theory *fits* the way that they use their sentences?
- Couldn't we make **any** theory of truth fit with the way that a speaker speaks?

LongLostish Again

- Here was our working theory for LongLostish
 - An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is white
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- According to this theory, 'Zagga gava gai' is true iff a white rabbit is nearby
- But suppose one day a *silver* rabbit comes by, and people shout 'Zagga gava gai!'
- How, if at all, should we revise our theory?

Option A

- One option would be to revise our theory as follows:
 - An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is white-or-silver
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- Now our theory implies that 'Zagga gava gai' is true iff a white-or-silver rabbit is nearby
- So the locals were saying something true when they shouted 'Zagga gava gai!' in the presence of a silver rabbit

Option B

- But another option would be to leave our theory exactly as it was:
 - An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is white
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- Now the locals have said something false when they shouted 'Zagga gava gai!' in the presence of a silver rabbit
- But we can explain *why* they said something false, by pointing out that it can be hard to tell white animals from silver animals in the wild, and so it is easy to get mixed up

Extending the Worry

- You might think that you could choose between these two options by checking how people use 'Zagga gava gai' in other contexts, and check whether they tend to use it just when white rabbits are nearby, or also when silver ones are
- Let's suppose that in general, the locals do say 'Zagga gava gai' when white rabbits are nearby and when silver rabbits are nearby
- We could still stick to our original theory of meaning if we insisted that when people said 'Zagga gava gai' when silver rabbits are nearby, they are making a mistake

Putting the Worry Starkly

- In fact, if we are willing to attribute the speakers of LongLostish enough false beliefs, we can make any theory of truth fit with the way they use their sentences
 - An object satisfies 'gava' iff it is a rabbit
 - An object satisfies 'bava' iff it is a monkey
 - An object satisfies 'zagga' iff it is electric pink
 - A sentence of the form ' α gai' is true in LongLostish iff something which satisfies α is nearby
- According to this theory, 'Zagga gava gai' is true iff an electric pink rabbit is nearby
- We can make this theory fit the fact that the locals say 'Zagga gava gai!' when white rabbits are nearby, simply by insisting that they falsely believe that the rabbits are pink!

Diagnosing the Problem

- All we can tell from the way that speakers use a sentence is whether they hold it to be true or hold it to be false
- But whether a speaker holds a sentence true depends on two factors:
 - (1) What the sentence means
 - (2) What the speaker believes
- If we knew what a speaker believed, then we could use the fact that they held a sentence to be true to figure out what the sentence means
- But in general, there is no way of figuring out what a speaker believes without already understanding what their sentences mean

Davidson's Solution: The Principle of Charity

- Davidson had a solution to this problem, which he called the **Principle of Charity**
- According to this principle, we should interpret other speakers in such a way as to **maximise the number of their true beliefs**
- In other words, our interpretation should make as many of the sentences which the speakers hold true actually come out as true
- That would stop us from using the theory which implies that 'Zagga gava gai' is true iff an electric pink rabbit is nearby
- That theory would require us to say that lots and lots of the sentences that LongLostish speakers hold to be true are actually false

Clarifying the Principle

- To be clear, when Davidson's Principle of Charity says that we should maximise the number of **true** beliefs that a speaker has, he means *true by our lights*
- That is not because Davidson is some kind of relativist about truth
- It is because we have no choice but to use our own best methods to figure out what is true
- So when **we** are trying to interpret someone else, we are trying to maximise the number of their beliefs which are true by **our** best methods

Clarifying the Principle

- To be clear, when the Principle of Charity says that we should **maximise** the number of true beliefs that a speaker has, it is not telling us to make **all** of their beliefs true
- It is fine to say that speakers make mistakes **sometimes**
- What the Principle requires is that **most** of the speaker's beliefs be counted as true, and any mistakes we attribute should be explicable
- The idea is that we can only understand a *disagreement* between what we believe and what the speakers believe if we place it against a background of broad *agreement*

The Status of the Principle

- Lastly, it is important to be clear about the *status* of the Principle of Charity
- Davidson does not think that our interpretations should maximise the number of true beliefs that a speaker has because we have some sort of guarantee that humans won't make too many mistakes
- Rather, his point is that if we ever come up with an interpretation which attributes lots of false beliefs to speakers, then that is automatically reason to think it isn't a very good interpretation
 - If you interpret me as saying something absolutely ridiculous, and cannot explain why I would think something so ridiculous, then that just goes to show that your interpretation was a bad one

Summing Up

- Davidson wants to use Tarski-style theories of truth as theories of meaning
- We can test whether a given theory of meaning is a good one by using it to make predictions about the truth-conditions of a range of sentences, and checking whether those predictions fit with the way that the sentences are used
- In performing this test, we have to employ the Principle of Charity, and insist that if the only way of making a given theory 'fit' with the way that people are speaking is by attributing them lots of false beliefs, then it isn't a good theory of meaning

Davidson and Quine

- It is easy to see the strong similarities between Davidson and Quine, but there are important differences too
- Davidson is not sceptical about the concept of meaning
- Davidson is not trying to reduce meaning to behaviour
- Davidson is trying to explain the relationship between meaning, truth and belief
- None of these concepts are to be reduced to anything more basic
- They are to be taken just as they are; the job of a philosopher is just to see how they link together

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

Objecting to Davidson's Programme

- Lots of philosophers have presented lots of objections to Davidson's programme
- We won't have time to go through them all here, but you can find many of them in the readings on the VLE Reading List
- In this part of the lecture, I want to quickly go through one objection
- This objection is really a straightforward misunderstanding, and it is helpful to see where it goes wrong

The Meaning of an English Sentence

- According to Davidson, we can give the meaning of 'Socrates is wise' like this:
 - (1) 'Socrates is wise' is true in English iff Socrates is wise
- The 'iff' is the material biconditional, meaning that (1) is true whenever the two sides have the same truth-value
- In that case, we can swap the 'Socrates is wise' on the right for any other true sentence we like, and the whole biconditional will stay true:
 - (2) 'Socrates is wise' is true in English iff Trump is president
- So isn't Davidson stuck saying that 'Socrates is wise' means that Trump is president?

The Mistake in this Objection

- The mistake is in thinking that we can see what 'Socrates is wise' means just by noting that the following is true:
 - (1) 'Socrates is wise' is true in English iff Socrates is wise
- Really, we give the meaning by showing how a sentence like (1) follows from a good theory of meaning

(A theory of meaning which, as a whole, fits well with the way that speakers use their sentences)
- But while a good theory of meaning for English would logically imply (1), there is no reason to think it would logically imply:
 - (2) 'Socrates is wise' is true in English iff Trump is president

The Mistake in this Objection

- In fact, there is every reason to think that a good theory of meaning **wouldn't** imply:
 - (2) 'Socrates is wise' is true in English iff Trump is president
- Recall that when we develop a theory of meaning, we need to assign semantic properties to the expressions which make a sentence up
- In order to get a theory of meaning to imply (2), we would need to base it on axioms like this:
 - 'Socrates' refers to Trump
 - Something satisfies 'x is wise' iff it is president
- But a theory built on those axioms would clearly fit our whole language very poorly

A Bad Theory of Meaning

- Suppose we tried to develop a theory of meaning for English with these axioms:
 - 'Socrates' refers to Trump
 - Something satisfies 'x is wise' iff it is president
- That theory would go on to tell us that 'Socrates is Greek' is true iff Trump satisfies 'x is Greek'
- Assuming that an object satisfies 'x is Greek' iff it is Greek, this gives us this truth-condition:
 - 'Socrates is Greek' is true in English iff Trump is Greek
- On this theory, then, 'Socrates is Greek' is false, even though we all think it is true
- Similar problems will ramify out, and so the Principle of Charity will tell us that this is a bad theory

A Good Theory of Meaning

- By contrast, things would go much better if we developed a theory of meaning for English with these axioms:
 - 'Socrates' refers to Socrates
 - Something satisfies 'x is wise' iff it is wise
 - Something satisfies 'x is Greek' iff it is Greek
- Then we can infer:
 - 'Socrates is wise' is true in English iff Socrates is wise
 - 'Socrates is Greek' is true in English iff Socrates is Greek
- Now we end up saying that 'Socrates is Greek' is true, just as we all believe it is

The Important Point

- Davidson **does not** think that you can figure out what 'Socrates is wise' means just by noting that the following is true:
 - (1) 'Socrates is wise' is true in English iff Socrates is wise
- You need to show how (1) follows from a theory of meaning which fits well with our general practice
- We can of course do that, but we couldn't do the same for:
 - (2) 'Socrates is wise' is true in English iff Trump is president
- A theory of meaning which implied (2) would have to be built on axioms about the meanings of words which fit poorly with the way we actually use sentences

Davidson's Theory of Meaning

Re-Cap: Grice's Theory of Meaning

What should a Theory of Meaning Look Like?

Theories of Truth

Radical Interpretation

The Principle of Charity

An Objection

Conclusion

The Search for a Theory of Meaning

- Davidson wants to figure out how to construct a **theory of meaning**
- A theory of meaning for a language L would somehow present the meaning of every sentence in L
- There are infinitely many sentences in L , so a theory of meaning would have to be **compositional**
- The theory would have to explain what a sentence means by showing how its meaning is determined by the meanings of the expressions which it is built out of

From Meaning to Truth

- Davidson thought that we could use a Tarskian theory of **truth** for L as a theory of **meaning** for L
- Tarskian theories of truth are **compositional**
 - The truth-conditions of whole sentences are determined by the semantic properties of the expressions they are built out of
- A good Tarskian theory of truth is meant to satisfy **Convention T**
 - A good theory of truth for L will yield outputs of the form 's is true in L iff p ', where s is a sentence from L , and ' p ' is the meta-linguistic translation of s

Testing a Theory of Truth

- However, while Davidson agreed that a good theory of truth should meet Convention T, he recognised that he could not use this as a test for whether the theory is good
- Convention T appeals to the concept of *translation*, which in turn is defined in terms of *meaning*, and Davidson wanted to co-opt a theory of truth as a theory of meaning
- Davidson thought we could see how we should test a theory of meaning by thinking about cases of **radical interpretation**

Radical Interpretation

- If we are trying to interpret a language we do not understand, we should start by watching how speakers use whole sentences
- We should then try to come up with truth-conditions for those sentences which fit the way that they are used
- After we have assigned truth-conditions to enough sentences, we can devise a theory which implies that those sentences have those truth-conditions
- That theory will have axioms which assign semantic roles to the expressions which make sentences up
- We can use those axioms to make new predictions about the truth-conditions of other sentences, and we can then check how well those predictions fit the way that the speakers use their sentences

Semantic Holism

- This way of thinking about interpretation leads to **semantic holism**
- According to semantic holism, it does not really make sense to ask what a given sentence means in isolation
- We can only ask what a single sentence means in the context of a whole language
- That's because the only way of investigating whether we have interpreted a given sentence correctly is by using that interpretation to make predictions about what **other** sentences mean, and then checking whether those predictions are right

The Principle of Charity

- A radical interpreter must follow the Principle of Charity
- According to that principle, when we interpret other people, we must do so in a way which maximises the number of true beliefs that they hold
 - We want a theory which, for the most part, makes a sentence true just when the speakers hold it to be true
- Without the Principle of Charity, **any** theory meaning could be made to fit with the way that speakers use the sentences in their language

Problems for Davidson

- Lots of philosophers have raised lots of objections to Davidson's programme
- First, it is not clear that you can really construct a Tarskian theory of truth for natural languages
 - Tarski designed his theories to work for **formal** languages, and **natural** languages use lots of devices that do not appear in formal systems
- And some philosophers have argued that even if you can cook up a theory of truth for a natural language, it wouldn't work as a theory of meaning
- We have only had the time to look at one objection today, but you can find more in the Reading List, and I would be happy to chat about them in my office hours

Applications Outside of the Philosophy of Language

- Despite the objections, many philosophers and linguists find Davidson's approach to meaning attractive
- Part of the attraction is due to the fact that Davidson's approach seems to have important consequences for issues **outside** of the philosophy of language
- I will end by quickly sketching one of those consequences, which you can think on later
- If you want to know more about these consequences, you should read these two papers by Davidson, both of which are on the Reading List:
 - The Method of Truth in Metaphysics
 - On the Very Idea of a Conceptual Scheme

Davidson against Scepticism

- According to Davidson, we have to use the Principle of Charity when interpreting speakers
- This applies to ourselves as much as it applies to anyone else
- If someone wants to interpret what we say, then they must start by assuming that most of our beliefs are true
- Davidson uses this point to argue against radical scepticism, according to which the vast majority of your beliefs are false

Davidson against Scepticism

- Imagine that there was an **omniscient interpreter**, who knows everything about the external world, and she wants to interpret you
- She will have to follow the Principle of Charity, and so interpret you in a way which makes most of your beliefs true by her lights
- However, since she is omniscient about the external world, if the majority of your beliefs are true by **her** lights, then most of your beliefs are **true**, full stop
- So any kind of radical scepticism, according to which most of your beliefs are false, is wrong
- Any interpretation which attributed you mostly false beliefs would just be a bad interpretation!

Tomorrow's Seminar

- The reading for tomorrow's seminar is:
 - Davidson, 'Radical Intepretation'
- Access to this paper can be found on the VLE Reading List

Next Week's Lecture and Seminar

- For next week's lecture, read:
 - Lycan, *Philosophy of Language*, Chapter 14
- For next week's seminar, read:
 - Davidson, 'What Metaphors Mean'
- Access to the the textbook chapter and the seminar reading can be found on the VLE Reading List