(1) One of the things we discussed before, when going through Frege's view of sense and reference, I mentioned that he proposes a *principle of compositionality*. This is a principle along the line os:

> The meaning of a sentence is determined by the meaning of its constituents and the way in which they are combined.

As I mentioned then, the reason why we might be interested in such a principle is that it helps us explain how we can learn a language which can produce an unlimited amount of understandable sentences even though we only have finite information available to learn the language from.

(2) To help us talk about Davidson's theory, I want to borrow a distinction from Michael Dummett (1991).

(a) A *semantic theory* for a language L is a specification of meaning for every sentence of L.

(b) A *theory of meaning* is a theory which describes how semantic theories are structured.

So, a theory of meaning is an entire conception of how meaning works, such as verificationism or propositional theories. What Lycan (2019) slightly confusingly calls a "theory of meaning for L" is exactly the same as what I call a semantic theory for L.

(3) Compositionality, according to Davidson, is one of the central features which allows us to explain linguistic communication. A semantic theory for a language L must be possible to generate from a finite collection of rules for how expressions can be combined.

(4) Davidson also demands that a semantic theory for L entails *truth conditions* for the sentences of L. When we know what a sentence means, then we must know under what conditions it would be true or false. In this way, Davidson differs from the verificationists by saying that to understand a sentence we only need to know what it would take for it to be true. We don't need to know what possible experience would show us that it is true.

(5) So, what Davidson wants is precisely the sort of theory which Tarski has provided for formal languages. A semantic theory for L is precisely a theory which entails *T-sentences* for all sentences in L.

(a) A T-sentence is, for example, that

> 'Snow is white' is true if and only if snow is white.

(b) For an arbitrary sentence S of L, its T-sentence is of the form

> 'S' is true if and only if P.

where P is a sentence which is true exactly when S is.

(c) Now, a semantic theory for L must be formulated in a language which can express statements, like T-sentences, which mention sentences of L. But to avoid paradoxes, like the

Liar, this language generally can't be L itself. For that reason we call the language in which the semantic theory for L is expressed a *metalanguage* for the *object language L*.

(d) So, in a T-sentence S is a sentence of L which has been quoted, and thus mentioned, into the metalanguage whereas P is a sentence of the metalanguage which is used as normal. You can think of a T-sentence as expressing an explanation for which sentence in the metalanguage shares truth conditions with the sentence S of the language L. For example,

'Snö är vitt' is true if and only if snow is white.

It works as a kind of translation which says that the quoted Swedish sentence is true exactly when the used English sentence is.

(6) Tarski also shows how we can extend this kind of theory to more complex sentences in a compositional way through what he calls the *convention T*.

(a) The idea is that you start with a grammar for how sentences can be constructed in the language. For every rule of grammar which allows us to construct a complex sentence from simpler ones, there is a semantic rule for how the truth conditions of the resulting complex is determined in terms of the truth conditions of the simpler constituents.

(b) Take, for example, the grammar of first-order logik. Then our basic constituents are predicates and singular terms. Our most basic rule of construction is that a complete sentence can be formed by applying a predicate to a singular term. Further, we can construct complex sentences by combining them through logical connectives.

(c) For this kind of language, convention T would look something as follows. I've not included the rule for every single connective mainly for reasons of space.

'Pt' is true if and only if what 't' refers to satisfies 'P'.

'$\neg\varphi$' is true if and only if 'φ' is not true.

'$\varphi \wedge \psi$' is true if and only if 'φ' is true and 'ψ' is true.

'$\exists x Px$' is true if and only if 'Pt' is true for some 't'.

The truth conditions on the left of the biconditional are defined in terms of the ones on the right, which are all truth conditions for simpler sentences. So, in this way we can reduce the truth conditions for complex sentences into combinations of truth conditions for simpler ones until we reach the basic sentences of the form Pt.

(d) Accordingly, all we need to generate T-sentences for any sentence in the object language is a specification of which predicates apply to which singular terms. That finite basis of information, together with this finite list of rules, is sufficient to explain the truth conditions for any of the infinitely many sentences of L.

(e) This is, according to Davidson, the general form of a semantic theory for a language. It specifies the basic symbols and provides a finite set of rules which define the truth conditions for complex sentences in terms of the truth conditions of their grammatical constituents. For more complicated languages, like English, we would need many more complicated rules to formulate convention T since it has a much more complicated grammatical structure than a first-order language. But the principle would be the same.

(7) Now, Lycan (2019) gives a somewhat misleading picture of Davidson's theory. What I've said so far fits with what he has to say and is something Davidson would commit to. But there are a few points which are crucial to Davidson which Lycan doesn't mention. We can see them in Quine's example on the indeterminacy of translation which Davidson takes very seriously.

(8) **(a)** Quine asks us to imagine that we are anthropologists who are contacting a new group of people. As such, we don't share a language with them.

(b) One thing we notice is that they seem to utter 'Gavagai' when in the presence of rabbits. Hence, we write down the hypothesis that it means rabbit.

(c) But it's really only our linguistic complacency which supports that conclusion. All our data fits equally well with the hypothesis that the word means 'non-detached rabbit-parts' or 'time-slice of a rabbit'.

(d) Just because *we* conceptualise rabbit as more simpler or more fundamental than the alternative concepts doesn't mean that the speaker we're interpreting does the same.

(e) For example, one day we might hear a speaker say 'Gavagai' and point to a hare instead. One option is to conclude that the word means 'rabbit or hare'. Another is that the speaker mistook a hare for a rabbit.

(9) Davidson thinks that Quine's argument is just as valid when we're trying to understand speakers of the same language as ourselves. We're engaging in what he calls *radical interpretation*.

(a) When we try to understand another speaker, what we're doing is trying to construct a semantic theory for their language. But, just like in Quine's example, it's impossible to determine which semantic theory is correct. There's no unique correct answer to which truth-conditions correspond to a sentence.

(b) We can't really distinguish between what beliefs we attribute to a speaker and how we interpret the meaning of their utterances. That is, we can't determine if a particular T-sentence for an utterance fits the data; there might be differences among the rest of their beliefs or meanings of other sentences which make a difference. A semantic theory for, an interpretation of, someone's language can only face experience as a whole.

(c) Further, in order to even try to interpret someone's language we have to assume that the majority of their beliefs are true. Otherwise we can't use their linguistic behaviour as data for what truth conditions they ascribe to their sentences. This is what Davidson calls the *principle of charity*.

So, Davidson dismisses the idea of truth conditions as meaning-entities. Unlike other truth-conditional views, his is not an entity theory. He also accepts that meaning is holistic even as he also thinks it can be analysed compositionally.

(10) Lycan also implies that Davidson accepts a correspondence theory of truth. He does not. On the contrary, he doesn't think that truth can be analysed. It's a primitive concept which we have to take for granted. Allt that there is to say about the concept of truth is that it abides by convention T, which relates ascription of truth in a metalanguage to the acceptance of sentence in the object language. Thus, he accepts the claims that

<p align="center">'Snow is white' if and only if snow is white.</p>

<div align="center">'Niklas' refers to Niklas.</div>

but is not committed to the view that they express any word-world relation.

(11) Many of the issues which Davidson's theory face have to do with the fact that T-sentences are entirely *extensional*. They can't distinguish concepts which just happen to apply to the same things or are true at the same time.

(a) Further, just like most other theories we've discussed, it only discusses descriptive fact-stating language.

(b) One specific problem has to do with how to handle indexicals, like 'I', or demonstratives, like 'that'. How do we give a T-sentence for someones utterance that

<div align="center">I am in Lund.</div>

<div align="center">That's Niklas coffee mug.</div>

when we're not the same person or in the same place as the original speaker?

(c) It happens to be the case that mammals with a heart also have kidneys. So, the sentence

<div align="center">'Dogs have hearts' is true if and only if dogs have kidneys.</div>

is a true T-sentence. But we wouldn't want to say that it explains the meaning of the claim that dogs have hearts.

(d) Similarly, Davidson has difficulties explaining modal claims or other connectives which aren't purely truth-conditional.

(12) Our second theory of the day is H.P. Grice's *psychological* theory of meaning. Just like Strawson thought that speakers rather than terms themselves are what refer, Grice thinks that speakers rather than words provide meaning. When linguistic expressions are abstracted away from their speakers, we make it difficult to see how what we choose to say relates to what we want to convey.

(13) If we think of the communicative roll of an expresion purely in terms of the information it conveys, it's not clear that we can make a distinction between linguistically communication content and information gained through reliable observation. Consider the situations described by the following sentences.

<div align="center">The spots mean that the patient has measles.</div>

<div align="center">The symbol means that you've passed the exam.</div>

Both situations are ones where we can acquire information from something we observe in our environment. But, according to Grice, only the latter can be considered a symbol with *communicative* content. Thus, it's only in the second sentence that 'means' should be understood as describing linguistic expression.

(14) Grice thinks that the important difference is that linguistic communication is something we do intentionally. And my intention when making an utterance plays an important role for the hearer to understand what was said. Linguistic utterances are a much richer source of information than mere reliable observation precisely because it occurs through an interaction between the participants *communicative intentions*.

(15) What makes Grice's theory psychological is just that intentions, which are mental states, play a central role. The idea here is to reduce meaning to psychological states, although in a much more sophisticated way than Locke tried to do.

(16) The first step to Grice's theory, then, is to give an analysis of what it means to say that

$$S \text{ meant } P \text{ with the utterance } x.$$

The explanation comes in three parts. The analysis as I formulate it here is limited to attempts to convey beliefs, but Grice is open to the idea that we can intend to convey other cognitive states.

   (a) S uttered x with the intention that the audience forms the belief P.

   (b) S intended that the audience recognises the intention described in (a).

   (c) S intended that the audience forms the belief that P by the means of recognising the intention described in (a).

(17) As this is an analysis of waht a speaker conveys with a specific utterance, Grice calls this *speaker-meaning*. As we can see, intentions aren't the only mental term used in the analysis of speaker-meaning; recognition and belief also play a crucial role.

(18) Next up is to try and eplain *sentence-meaning*, that is what the sentences themselves mean, in terms of speaker-meaning.

   (a) He begins by describing how an unstructured expression can acquire conventiional meaning from the speaker-meaning it's usually uttered to convey. An unstructured expression is one which can't be compositionally analysed into simpler constituents.

   (b) The idea is that speakers can have the habit to use an expression x to speaker-mean P. Grice says that they have the procedure to use x to speaker-mean P in their repertoire.

   (c) So, Grice says that an unstructured expression x means P for a speaker-group G if and only if

      (i) Many of the speakers in G has the procedure to use x to speaker-mean P in their repertoire.

      (ii) The speakers maintain this procedure because other speakers in G have the same procedure in their repertoire.

   (d) In this way, there is a shared and conscious practice of using x to speaker-mean preciselt P in the group.

(19) But explaining how unstructured expression can have conventional meaning isn't sufficient. If the principle of compositionality is right, then complete sentences are structured and have their meaning determined by the meaning of their constituents. Further, this method doesn't explain how new and unfamiliar sentences can have sentence-meaning. Precisely because they are novel, the members of the speaker-group can't have a shared procedure of using them to speaker-mean anything in particular.

(20) The idea when it comes to structured expressions is to use the above method on their unstructured constituents. Take the following sentence as an example.

<center>Niklas doesn't like ice cream so Niklas isn't nice.</center>

We can decompose this complex sentence into two parts.

<center>Niklas doesn't like ice cream.</center>

<center>Niklas isn't nice.</center>

which have been connected through a certain grammatical construction, in this case a conditional.

(a) We can consider these two subsentences as unstructured expressions and apply the above method to explain their sentence-meanings in terms of what procedure to use them which the speakers have in their repertoire.

(b) If we also have a procedure in our repertoire to compose basic procedures to resultant procedures according to the grammar of the structured expression, then we can say that the resultant procedure specifies the sentence-meaning of the complex sentence.

(c) So, the sentence-meaning for the complete sentence is determined by the procedure which results when we compose the procedures for 'Niklas doesn't like ice cream' and 'Niklas isn't nice' according to them being premise and conclusion for a conditional statement.

(d) Hence, we can explain compositionality and thus sentence-meaning novel expressions in terms of the sentence-meaning of their unstructured components which are already in use.

(21) But this explanation also doesn't quite go all the way. Since Grice has only analysed speaker-meaning for complete sentences, the kind which can express a belief P, sentences are the smallest unstructured expressions available to this explanation. Individual words, terms, or predicates cant be ascribed a speaker-meaning this way and, hence, can't acquire expression-meaning. So, to fully explain compositionality and unfamiliar expressions, the theory would need to explain how individual words contribute to the speaker-meaning of a complete sentence.

(22) There are also some classic objections to Grice's analysis of speaker-meaning.

(a) Philosopher's often talk to themselves when thinking through an argument. You might even utter the very same sentence several times to consider it carefully. But there's no hearer present and no communicative intention which needs to be the same between the subsequent utterances of the sentence to ensure that the same content has been speaker-meant. Nonetheless, we want to say that something was said and that all the utterances of the sentence meant the same thing.

(b) When I tell you that

<center>Grice advocated a psychological theory of meaning.</center>

I can't intend for you to come to believe it. Since you're good students, I already know that you've read the chapter on Grice before the lecture. And if I present some argument with the intention that you believe it's conclusion, I don't want you to believe it *because that was my intent*. I want you to believe it because the argument is good.

(c) The connection between speaker-meaning and the words we choose to use it seems far too loose. Searle gives the example of an American officer captured by Italian soldiers during World War II. In the hope that they also don't know German, he tries to trick them into setting him free by shouting the only German sentence he knows in a commanding voice.

<center>6</center>

Kennst du das Land wo die Zitronen blühen?

> When he utters this sentence he intends for the hearers to form the belief that he is German, that they recognise this intention, and by recognising this intention forming that belief. But it doesn't seem right to say that his utterance of this sentence *means* that he's German.

(23) So, Grice has a reverse strategy from the one we've seen invoked a few times. Many theories start with an analysis of sentence-meaning or semantic-meaning, which is independent of speakers and intentions, and then tries to explain speaker-meaning as a pragmatic consequence of it. Grice, on the other hand, starts with what speakers intend by their uses of sentences and tries to explain sentence-meaning in terms of that.