

Meaning and Use

Robert van Rooij

1 Introduction

This paper deals with the meaning of natural language expressions, and how meanings of expressions are used in communication. The two disciplines that talk most about meanings of expressions are linguistics (semantics and pragmatics) and philosophy. This paper is about topics discussed in both disciplines. The first part of the paper is more philosophical in nature and discusses what is meaning in the first place, and how it is related with reference. The second part is concerned with the relation between semantics and pragmatics.

Although certainly not uncontroversial, there can be little doubt that what is known as ‘formal semantics’ is the most productive brand of natural language semantics within linguistics (cf. the popular introductions, Chierchia & McConnell-Ginet (1990) and Heim & Kratzer (1998)). In formal semantics, the meaning of a sentence is its truth conditions. Once we adopt a truth conditional concept of meaning, it is natural to think of reference as depending on meaning and that semantics is consistent with the Chomskyan cognitive, and individualistic, program in linguistics. In the first part of this paper we will first discuss two well-known problems for a particular way to combine these ideas: ‘Putnam’s paradox’ and Kripke’s counterexamples. Next, we discuss how a causal theory of reference can overcome these problems when framed within a two-dimensional conception of meaning, and what this means for the interpretation of the latter framework.

The second part of the paper investigates the relation between semantics and pragmatics. First it discusses that what is communicated with the use of a sentence on top of its semantic meaning: Gricean *conversational implicatures*. The discussion will be limited to implicatures generated by Grice’s maxim of Quality and his first submaxim of Quantity. *Speech acts* will be

discussed afterwards. First *assertions*, and the idea to think of a *presupposition* of a sentence as a felicity condition for the appropriate use of the sentence. Then *questions*, focussing on the issue whether Searle (1969) was right in claiming that the meaning of the interrogative sentence ‘Is the door open?’ is the same as that of its declarative analogue ‘The door is open’, the difference being just the way (speech act) in which the sentence is used. Finally we will deal with *imperatives* and *permissions*. In this part it will be discussed whether a performative or an assertive analysis of disjunctive permission sentences is most suitable to account for their free choice inferences.

2 Meaning and reference

2.1 Meaning determines reference

The perhaps most ‘natural’ conception of ‘meaning’, at least in its point of departure, identifies ‘meaning’ with *naming*. The meaning of an expression is that what the expression *refers to*, or *is about*. What meaning does is to establish a *correspondence* between expressions in a *language* and things in the (model of the) *world*. For simple expressions, this view of meaning is natural and simple. The meaning of a proper name like ‘John’ or definite description like ‘the number of major planets’, for instance, is the object or number denoted by it, while the meaning of a simple declarative sentence like ‘John came’ could then be the *fact* that John came. Beyond this point of departure, things are perhaps less natural. What, for example, should be the things out in the world that common nouns and a negated sentence like ‘John didn’t come’ are about? One can, of course, assume that they refer to real existing properties and negative facts, but these assumptions make our initial hypothesis immediately less appealing. Except for conceptual worries, this referential theory of meaning gives rise to a serious empirical difficulty as well: the *substitution problem*. Assuming, by the principle of compositionality, that the meaning of a complex sentence depends only on the meanings of its part and the way these parts are put together, it follows that if two expressions have the same meaning, one can substitute the one expression for the other in a complex sentence without change of meaning. But because there are 9 major planets in our solar system, on the theory of meaning at hand the expressions ‘9’ and ‘the number of major planets’ refer to the same thing, and thus have the same meaning. Still, we cannot

substitute the expressions ‘the number of major planets’ for the number 9 in the sentence ‘It is necessary that 9 is bigger than 7’ without changing its truth value. The natural conclusion is that the meaning of an expression like ‘the number of planets’ should not be identified with its referent.

It seems natural to assume that to be a competent speaker of English one has to know what it means for ‘John came’ to be true or false. So a minimal requirement for any theory of meaning seems to be that one knows the meaning of a declarative sentence if one knows under which circumstances it is, or would be, true. The proposal of formal semanticists to solve our above conceptual problems is to stick to this minimal requirement: identify the meaning of a declarative sentence with the *conditions*, or *circumstances* under which the sentence is *true*. These circumstances can, in turn, be thought of as the ways the world might have been, first order models, or *possible worlds*. Thus, the meaning of a sentence can be thought of as the set of models, or possible worlds, in which it is true. This latter set is known in possible worlds semantics as the *proposition* expressed by the sentence.

Possible world semantics can account for the substitution puzzle when it assumes that the different expressions have different referents in at least some possible worlds. The most natural way in which this can be accounted for is to follow Frege (1898) and make a distinction between the *meaning* and *reference*, or denotation, of an expression. In possible world semantics this distinction can be modeled by assuming that the denotation of an expression in a possible world is simply an object, and that its meaning is a non-constant *function* from possible worlds to its denotation in that world. This is natural for definite descriptions like ‘the number of major planets’, but what about other types of expressions for which such substitution puzzles can arise, like proper names (‘Hesperus is Phosphorus’), and common nouns (‘Water is H_2O ’)?

Well, it seems that with the use of a proper name, or common noun, we associate a cluster of predicates or properties. The natural suggestion then is to *define* the meanings of these expressions in terms of these predicates or properties: they (or a sufficient number of them) give the list of *necessary* and *sufficient* conditions that an object, or stuff, has to satisfy in order to be denoted by the proper name or common noun.¹ Obviously, the above substi-

¹This description theory of reference is only one example of a cluster theory of meaning. Many scholars, ranging from philosophers like Wittgenstein (1953) and Searle (1958), linguists like Katz & Postal (1964), and psychologists like Rosch (1978), have proposed that it are clusters of characteristic properties that (help to) identify the denotation of an

tution puzzles for proper names and common nouns do not arise on such a view. For instance, although the names ‘Hesperus’ and ‘Phosphorus’, or ‘Cicero’ and ‘Tully’, actually refer to the same individual, one can propose that they don’t have the same meaning, and thus cannot always be substituted for each other in a sentence without change in meaning.

The truth-conditional tradition in semantics has its source in the work of logicians and philosophers like Frege who held a rather anti-psychologistic view towards meanings. However, once we adopt the above possible world semantics and think primarily of truth-conditions rather than of truth, and of reference as depending on meaning, we can think of semantics as being consistent with the Chomskyan cognitive, and individualistic, program in linguistics. Specifying the meaning of an expression should be consistent with knowing the meaning of that expression, and this, of course, is true if one takes the meaning of a proper name, or common noun, to be the set of properties associated with that expression. Moreover, the truth-conditional view on sentence meaning seems natural as well, because from a cognitive perspective this means that knowing the meaning of a sentence is to know under which circumstances it is true. In this way model theoretic, or possible world, semantics can account for the primary task of natural language semantics, at least when seen from a Chomskyan perspective: to account for pretheoretical judgements of speakers concerning semantic relatedness of expressions of a particular language, in particular of the relation of entailment.

More in general, specifying meanings as in standard possible world semantics is consistent with the computational model of the mind which has become fundamental for cognitive science. The view that the meaning of an expression, or of an internal state, determines, but is independent of, what that expression or state is about, is compatible with the computational model of the mind which sees interpretation and the explanation of behavior as involving only internal states, or of internal models of external states of affairs. It is true that according to some cognitive scientists meanings just *are* internal states, and why should model theory be of relevance here? But, then, the computational model of the mind favors a *functional* view of internal states: it is not an internal state all by itself that has a meaning, but rather the (abstract) function that state has to explain an individual’s overall behavior. Model theory is well suited to account for such abstract functions.

expression in a particular context of use, or what a thought is about.

2.2 Problems for the standard conception of meaning

However appealing and natural this combination of possible world semantics and the cluster theory of reference might be, it gives rise to at least two problems, one conceptual and one empirical in nature. In the following subsection I will discuss both problems, and some suggested solutions of them. As we will see, both problems are, in fact, independent of possible world semantics as such, and concern only the cluster theory of reference.²

2.2.1 Intended interpretation and meaning holism

A first problem concerns the predicates, or properties, used in the description that is supposed to identify the referent. According to the cluster theory of reference, a speaker refers with ‘*N*’ to *a* because *a* is the unique individual or stuff that satisfies the set of predicates that the agent associates with ‘*N*’. We can think of this set of predicates as the speaker’s representation of *a*. So, this analysis explains the speaker’s reference of *a* by ‘*N*’ in terms of the reference of the predicates associated with ‘*N*’. But that only gives rise to the questions *what* those predicates themselves refer to, and *why* that they do so. The standard cluster theory of meaning doesn’t seem to do more than explaining one part of the language in terms of other parts – the terms in which the descriptions are given. Obviously, our problem of why one type of expression refers to what they do is not really solved, but only replaced by the same problem for another type of expression. But perhaps we should not think of these other type of expressions as belonging to the same external language; perhaps they are expressions of an *internal* language, or some other kind of internal representations of an agent. It all doesn’t matter much: in whatever way we represent the speaker’s meaning of ‘*N*’ in terms of a set of internal representations of general terms, it always gives rise to the further question of why these internal representations of general terms are about what they are in fact about.

One way to get out of our above regress problem is to propose that we can’t interpret the terms of a language individually, but that we have to do so simultaneously for the language *as a whole*. The idea would be that the terms refer to whatever things, properties, and relations that do the best job of making the set of sentences true that speakers in fact consider to be true.

²This is not to say that possible world semantics by itself doesn’t give rise to conceptual and/or empirical problems. It certainly does, but I will ignore those problems in this paper.

Unfortunately, Putnam (1981) has showed that this picture as such is not constrained enough to fix the meaning of the expressions of a language in the intuitively correct way. He elaborates on the model theoretic fact that for any consistent set of sentences a model can be constructed with a domain of individuals that is not of the wrong size: we can always come up with different sets of objects or different interpretation functions that make the same set of sentences true. From this fact, Quine (1960) concluded to the indeterminacy of reference: knowing the truth value of a collection of sentences doesn't mean that you know the references of its constituents. Putnam (1981) generalized this argument to intensional languages: even if one knows the truth value of a sentence in every possible world (its intension), this doesn't necessarily mean that one knows the intuitively correct intensions of its constituents. For instance, it is possible to formulate highly counterintuitive intensions for expressions like *cat* and *mat*, so that in the actual world they refer to trees and cherries, respectively, without affecting the intension of *The cat is on the mat*. To determine the meaning of the terms of our language, knowing the truth value or intension of a collection of sentences is not enough, because the terms of the language can be assigned weird and 'unintended' interpretations.

Though Lakoff (1987) has argued otherwise, this argument obviously does not rule out model theoretic linguistic semantics as such. In practice, it doesn't even seem to be of any importance for model theoretic linguistic semantics at all: it typically doesn't care much about the meaning of basic terms, and certainly not about why these expressions have the meaning they have. Natural language semanticists that use model theory to account for the meaning of expressions are interested only in how the meanings of more complex expressions can be explained in terms of the meanings of simple terms and some expressions with a logically fixed meaning. Still, if we believe that the world is as it is independently of our conceptions of it, Putnam's argument gives rise to two general questions: first, why are only some of all possible interpretation functions compatible with the way English is spoken; second, also why are only some of all possible interpretation functions compatible with the way basic common nouns are used in *any* possible natural language? The first of these questions asks why English expressions have the meanings they actually have, and how the meaning of a proper name and common noun is determined in the first place, while the second one asks for more general constraints on how meanings could be assigned to expressions. Both of these problems can be solved if we are able to supplement model theoretic semantics with natural constraints on reference. But where could

those constraints come from?

Perhaps we should not limit ourselves to behavior that involves verification or falsification of sentences, but should consider behavior in general, and how this is related with the beliefs and desires of the agents. We might then propose that meanings are assigned primarily to attitudes of agents, and that such an attitude is about, or directed to, an object, stuff, or state of affairs because the agent is disposed to perform actions that involve this object, stuff, or state of affairs. Unfortunately, assignment of beliefs and desires such that it fits, or explains, the behavior of agents won't be enough:

What makes an assignment of a system of belief and desire to a subject correct cannot just be that his behaviour and behavioural dispositions fit it by serving the assigned desire according to the assigned beliefs. The problem is that fit is too easy. The same behaviour that fits a decent, reasonable system of belief and desire also will serve countless very peculiar systems. Start with a reasonable system, the one that is in fact correct; twist the system of belief so that the subject's alleged class of doxastic alternatives is some gruesome gerrymander; twist the system of desire in a countervailing way; and the subject's behaviour will fit the perverse and incorrect assignment exactly as well as it fits the reasonable and correct one. Thus constitutive principles of fit which impute a measure of instrumental rationality leave the content of belief radically underdetermined. (Lewis, 1986, p. 38)

At this point Quine's or Davidson's principle of *charity*, or of *humanity* seems a natural extra constraint. This principle demands that we should not attribute too much irrationality to a person in order to explain his behavior. Lewis (1984) argues that making use of such a principle involves making additional constraints on what the meanings and/or references of expressions and internal states could be. He proposes³ that the intended interpretation function of our language, or of any natural language, is not as free as Putnam presupposes, because the meaning, or intension, of simple lexicalized predicates like 'cat' and 'mat' must refer to 'well-behaving' or 'natural' properties, and he proposes some constraints (mostly involving a notion of similarity) on what such natural properties and relations could be. Lewis suggests that when we limit ourselves to interpretation functions that map

³See also Gärdenfors (2000) for a similar proposal to solve Putnam's paradox.

the simple predicates we use to ‘natural’ properties, there is no longer any guarantee that (almost) any world, or model of it, can satisfy (almost) any collection of sentences, and thus meaning indeterminacy might be tackled. Although I believe that Lewis’ proposal makes sense (if interpreted without his realist’ baggage) to answer the second of our above problems, I cannot see how it could account for the intended interpretation of *English*. Lewis’ proposal still seems to leave open too many interpretation functions.

Before we will discuss another possible way to solve Putnam’s paradox, let us first discuss a second problem for the cluster theory of reference.

2.2.2 Empirical problems

The second problem for the cluster theory of reference is *empirical* in nature. Donnellan (1970) and Kripke (1972) have convincingly argued that this theory of reference leads to counterintuitive results for proper names. More in particular, they have shown that speakers can refer, and even can *intent* to refer, to particular individuals without being able to describe or identify those individuals. First, speakers can successfully refer to a particular individual without having a uniquely identifying set of descriptions in mind. Second, even if they have such a description in mind, they sometimes still refer to an individual that doesn’t satisfy this description.

By very much the same kind of arguments, Kripke (1972) and Putnam (1975) have convincingly argued that the set of properties that speakers or agents associate with *natural kind terms* should also not be equated with the meaning of the noun. This is made very clear by the ‘Twin Earth’ stories given by Putnam (1975) and others. These stories always involve a comparison between two almost identical persons (twins): one in the actual world and one in a counterfactual world, Twin Earth, minimally different from the actual world. In Putnam’s story, the stuff that the inhabitants of the counterfactual situation call *water* is superficially the same as the stuff *we* call *water*, but its chemical structure is not H_2O , but XYZ . If, then, both the earthling and his twin assert ‘Water is the best drink for quenching thirst’, intuitively they have said something different. But how can this be if they associate exactly the same description with the word and if speaker’s description determines reference? A similar ‘Twin Earth’ story invented by Burge (1979) shows that the problem is not limited to a small set of terms. In fact, stories can be invented for almost any expression to show that it is not the description that the speaker associates with an expression that determines

its extension. The reason is that the linguistic practices of members of the agent's community are crucial in determining the extension of a term.

Perhaps what counts, then, is not so much the descriptions the *speaker*, or the *relevant agent*, associates with it, but rather the set of descriptions that *most people*, or *the specialists in the relevant linguistic community* associate with it. It is then this set of descriptions that determines the reference. However, Donnellan and Kripke have argued that this, too, give rise to counterintuitive predictions for proper names, while Putnam (1975) shows the same for natural kind terms. Putnam's demonstration involves the same 'Twin Earth' story, but now set in 1750. Specialists on Earth and Twin Earth are not yet able to see any difference between H_2O and XYZ . But intuitively, even if a typical Twin-Earthian (twin-) English speaker utters 'Water is the best drink for quenching thirst' on Earth, he is not talking about H_2O .

2.3 The causal theory of reference

Kripke and Putnam claim that the meaning of at least proper names and natural kind terms is not the set of descriptions associated with them, but simply what they refer to. But this gives rise to the question of *why* these expressions have the references they in fact have. At this point, Kripke proposed his causal theory of reference. Kripke (1972) argues that proper name 'N' can refer to a , only if, and because, a is the entity that is the *source* of the reference-preserving link from the initial baptism of the expression to the speaker's use of the name. Evans (1973) was perhaps the first to propose that the causal theory of reference should be based on a causal theory of belief, or of information. He argued with Kripke that a causal link for proper names is necessary, but that this causal link should not be between the initial naming and the speaker's current use of the name, but rather between the *body of information*, or superficial properties, relevant to the speaker's use of the proper name on a particular occasion and the object that is the dominant causal origin or source of this body of information. An object can be the dominant source of a particular body of information even if it does not fit this information very well. It follows that if P is one of the properties we associate with 'N', we still do not know that the sentence 'N is P' is true by necessity. This causal theory of aboutness can also explain why Oscar, but not his twin, talks or has beliefs *about* H_2O if he uses, or considers, the term 'water' in Putnam's (1975) Twin Earth story.

The causal theory of reference, or of meaning, seems also the natural candidate to limit the possible interpretations of the expressions of ‘our’ language, or of our thoughts, so as to solve Putnam’s paradox.⁴

The causal account of meaning is not without problems. It is not clear how to cash out the causal account in a completely naturalistic way and there are problems of how to account for our intuitions that we can have false beliefs.⁵ Moreover, it is unclear how a causal theory could ever determine the meaning of functional words, or of prepositions like ‘in’. But it seems that the causal account of content leads to unsolvable problems even if the above problems can be accounted for. Once we accept that the content of an intentional state or expression is just the causal source of the state, or of the use of the expression, we are confronted again with many old problems. If the meanings of ‘Hesperus’ and ‘Phosphorus’ are just their referents, the substitution puzzle arises again: ‘Hesperus is Phosphorus’ is predicted to express the necessary true proposition. But, then, how can we account for the fact that agents seriously doubt that such statements are true? Thus, the causal theory seems to predict a notion of content that is sometimes *not fine-grained enough* to account for our intuitions. At other times, however, the causal account of content seems to predict a notion of content that is *too fine-grained*, or *too specific*. For instance, it seems to predict that attitude ascriptions can no longer do the job commonsense psychology tells us they do. A common sense explanation of why the Earthling and his counterpart drink so much of the stuff that in their respective communities is called ‘water’ if they are thirsty is that they think that what they call ‘water’ is the best drink for quenching thirst. The problem is that according to the causal conception of content it seems that the belief attribution ‘Oscar believes that water is the

⁴Putnam (1981) claims that making use of this causal story is just adding more statements to our consistent set of sentences. But then it doesn’t solve the problem because the predicates of these additional sentences might be interpreted in unintended ways as well. But with Lewis (1984) I think that this is only the case if one thinks of the causal theory in a ‘descriptive’ way, as a set of sentences that has to be made true by the interpretation function. But the causal story is not intended to be incorporated within the semantic content of what is said with the name, it rather determines the content itself from a more external point of view.

⁵One way solve both of these problems involves making use of so-called ‘normality conditions’. But in order for the resulting analysis to be wholly naturalistic, we need a naturalistic analysis of such conditions. A natural candidate that suggests itself to provide such an analysis is Millikan’s (1984) biosemantics. I am not sure, though, whether this theory can do the full job.

best drink for quenching thirst' is more specific than we want, because we know that Oscar cannot distinguish H_2O from XYZ . This problem is also of relevance to linguistics, because on the causal story Oscar doesn't even know what he himself is talking about when he is using the term 'water'. This seems to be incompatible with Chomskyan linguistics.

2.4 Two-dimensional semantics

It is an obvious observation that what is expressed by a sentence is *context-dependent*: in different contexts the same sentence can express different propositions. For instance, the proposition expressed by 'I am living in Amsterdam' depends on who is the speaker in that context. In Kaplan's (1989) theory of context dependence, contexts consists of certain aspects of a world, like speaker, hearer, time, etc, and sometimes also the world itself. A context partially determines *what* is said by a sentence, and this is still modeled by a set of possible worlds.

Kaplan's theory of context dependence can explain why there are two ways people can disagree about the truth value of a statement. Suppose that the speaker claims something by uttering a sentence, and the hearer disagrees. They can disagree because the hearer has *misunderstood* the speaker. The hearer has made a wrong guess about the context of utterance the speaker was in, and thus about the context-dependent proposition expressed by the speaker. It is also possible that they agree about what is said, but *disagree about the facts that determine the truth value* of what is said.

If both context and possible world are relevant for determining the truth value of a sentence, we might say that the meaning of a sentence is a relation between them, a two-dimensional intension. Following Kaplan, we can call this kind of meaning the *character* of a sentence. The character of a sentence is compositionally determined by the characters of its parts. If E is an expression, we might call $[E]$ the character of E . Given a context, c , $[E](c)$ is the *content* or *intension* of E . $[E](c)(w)$, finally, is the *extension* of E , if w is a possible world. The content of a sentence is a proposition, and its extension a truth value.

Kaplan's theory of context dependence allows us to distinguish different reasons why a sentence is 'necessary' true. First, what a sentence expresses in context c can be true in every relevant world, $[A](c) = K$, where K is the set of all relevant worlds. Sentences like 'Hesperus is Phosphorus' and 'I am John' used by John are necessary in this way, because the contents,

or intentions, of proper names and indexicals are constant functions. But it might also be the case that a sentence is true in every context in which it is expressed. If $w(c)$ gives us the world of c , this means that for all $c : w(c) \in [A](c)$ holds. For instance, an English sentence like ‘I am here now’ is necessary true for this reason. We can think of the set of contexts in which a sentence is true as a semantic object as well, and we might call it *the diagonal*.⁶ Important about this diagonal is that if a sentence contains a context dependent expression, it might be that the sentence expresses a necessary truth, although its diagonal doesn’t contain of all contexts.

Consider John’s uttering of ‘I am John’, for instance. We have seen that this sentence is necessary true – i.e., its content is the set of all worlds –, because both noun phrases refer to the same individual. Still, the sentence can, intuitively, be informative, because the hearer might be ignorant of the identity of the speaker, or at least doesn’t know that he is called ‘John’. This intuition can be accounted for within two-dimensional semantics by making use of the diagonal: the diagonal consists of some, but not all, contexts, because the hearer is unsure whether the actual context is one where the speaker is called ‘John’.

The examples discussed so far are rather straightforward, and involve all obviously context dependent expressions. But the two-dimensional analysis has been used to account for the other problems as well: it has been used to account for the fact that people can doubt whether the identity statement ‘Hesperus is Phosphorus’ is really true, and to explain why the belief attribution ‘Oscar believes that water is the best drink for quenching thirst’ is intuitively true, although Oscar cannot distinguish ‘real’ water, i.e, H_2O , from XYZ . The reason why this can be done is that not only the reference of expressions like ‘I’ and ‘you’ depend on contingent features of the context, but this is also true – at least according to the causal theory of reference – for proper names, natural kind terms, and, if we may believe Burge, in fact, for any other type of expression. But how would this go? Can we assume that the reference of a proper name is world-dependent, but not just because of the fact that objects could have been called differently?

The causal theory predicts that, in a sense, statements like ‘Hesperus is Phosphorus’ indeed only say something about the semantic rules of English. Still, it predicts that we can learn something non-linguistic if we are informed that Hesperus is Phosphorus. This is the case because even if the exact ref-

⁶After Stalnaker (1978).

erent of an expression used in a conversation is not clear, we normally *do* have a pretty good idea about what properties the referents of terms being used have. Thus, if we receive the information that the sentence ‘Hesperus is Phosphorus’ is true, we learn not only some facts about the semantics of English, but also some astronomical facts. We learn that the most salient heavenly body seen in the morning sky is identical with the most salient heavenly body seen in the evening sky, because we already believe and presuppose that we are in a world in which the referents of the relevant expressions have those properties. The same is true for a belief attribution like ‘The Babylonians didn’t believe that Hesperus is Phosphorus’. This sentence can be used to attribute a belief about a (partly) astronomical fact to the Babylonians, because we know all too well what information the Babylonians associated with the expressions.⁷

Notice, though, that there is a difference between the sense in which the reference of these expressions depends on context. The expression ‘I’ is context dependent, because *in English*, ‘I’ always refers to the speaker, and the same expression *of English* might be uttered by different speakers. The reference of ‘Phosphorus’ and ‘water’, on the other hand, are context dependent only because in different worlds they have a different meaning, or causal origin. But, of course, in that sense the meaning of ‘I’ is context dependent as well: in another world it might be that the meanings of the pronouns ‘I’ and ‘you’ are interchanged. If we assume that a language, or grammar, determines both what are expressions of a language and what these expressions mean, then it will follow that when the same expression has a different meaning (and not just reference) as it has in the actual world, the one who utters that expression in that other world uses a different language. But the same would be true when we consider proper names and natural kind terms: if we assume that the meanings (and not just references) of proper names and natural kind terms depend on their causal origin, we have to conclude that an expression used in a world with a different causal origin as in our world is

⁷The two-dimensional framework has also been used to explain what Oscar and his twin have in common when they say to themselves ‘Water is the best drink for quenching thirst’. Perhaps this is possible, but certainly not on what might seem to be the most straightforward way. The most straightforward way has it that the meaning of a common noun like ‘water’ is just a function from worlds to a particular stuff, such that in Earth this stuff is H_2O and in twin Earth it is XYZ . But, of course, this specification alone leaves open many many functions, and the vast majority of those functions give rise to completely unintended denotations. See the last paragraph of this section for more discussion.

part of a different language. Assuming that we speak a particular language, it follows that we sometimes don't know the meanings of the expressions we use. Though this might feel to some like a contradiction in terms, others will take this conclusion to be equally innocent as a philosopher's worry whether the tree in front of him is a 'real' tree, and not just a holographic image of it: only in case the language user or philosopher is contemplating such skeptical thoughts it has any practical consequence.

On the emerging picture, we cannot simply think of 'Phosphorus' as an expression of a language that might have different causal origins in different worlds.⁸ But why not?, you might wonder. Why not just think of the meaning of 'Phosphorus' as something like 'whatever is the causal origin of our use of the expression 'Phosphorus''? One reason is that on this view the causal theory of reference cannot by itself solve Putnam's paradox. We have explained the meaning of one expression in terms of the meaning of others and have thereby turned the causal theory of reference into a description theory that involves causal talk that itself might be interpreted in unintended ways. The basic point is that on this new analysis we still cannot put sufficient constraints on how to interpret expressions. Think of the analogy with indexical pronouns again. Even if we already know somehow that 'I' is an indexical pronoun, the meaning of 'I' would under a similar view not be much more than 'depending on the actual convention of the language, 'I' refers to either the speaker or the hearer'.⁹ It is even worse for expressions of whom a speaker doesn't know its type, or of an expression whose reference is not determined by its causal history, but – if we may believe Burge (1979) – still depends on external factors. For such type of expressions the meaning doesn't seem to be any more specific than, 'whatever this expressions means', which basically comes down to the view that the meaning of an expression is nothing but the expression itself, or the internal representation associated with it.

2.5 Vagueness and context dependence

Truth conditional semantics assumes that the meaning of a sentence is given by its truth conditions. The phenomenon of vagueness is a potential threat to this framework. Consider the sentence *John is tall* uttered in a situation

⁸See especially Stalnaker (1997) for a defense of this view.

⁹See especially Stalnaker (2001) for this type of argument.

where John is 1.80 meters tall. Is this sentence true or false in this situation? This is hard to tell.

Vagueness is standardly defined as the possession of borderline cases. The borderline cases of *tall* are normally said to be those individuals of which we cannot really say whether they are tall or not: a man who is 1.80 meters in height is neither clearly tall nor clearly non-tall. In three-valued logics one can handle this phenomenon by saying that such a man is neither in the positive extension of *tall*, nor in its negative extension. These positive and negative extensions are given by a partial valuation function. If John is a man who is 1.80 meters high, John falls in the *gap* of the positive and negative extension of *tall* and the sentence *John is tall* is predicted to be neither true nor false. A well-known problem of this analysis is that too many sentences are predicted to be of this fate: Both *John is tall or John is not tall* and *John is tall and John is not tall* are predicted to be neither true nor false as well, although the former, and certainly the latter, intuitively have a classical truth value: true and false, respectively. To get rid of this problem, Fine (1975) and Kamp (1975) proposed to make use of, next to a partial valuation function, also a set of *total* valuation functions. A total valuation function doesn't allow for gaps: each individual is either tall or not tall. Thus, a total valuation can make a partial valuation function *more precise*. If *John is tall* is neither true nor false according to the original partial valuation function, it will be either true or false according to each (accessible) total valuation function, because such total valuation functions have a specific cut-off point, or *delineation*, from whereon a man is counted as tall. However, different valuation function will have different such cut-off points. Crucial in supervaluation theory is the notion of *supertruth*: a sentence is supertrue iff it is true according to *all* (accessible) total valuation functions of the given partial valuation function. Similarly for the notion of *superfalsity*. Much of the appeal of supervaluations is that *John is tall or John is not tall* is predicted to be supertrue: although some total valuations count the former disjunct true and others the latter, they each make one of them true. Similarly for *John is tall and John is not tall*, which comes out as superfalsity. Supervaluation theory makes a difference between *local* and *global* notions of validity, defined in terms of the notions *truth* and *supertruth*. Just as it is supertruth that behaves classical, so it is for the *global* notion of validity. ϕ *superentails* ψ iff for all models M and partial valuation functions s , if ϕ is supertrue in M and s , ψ should be supertrue in M and s as well.

Supervaluation theory assumes that the partial valuation function with

which we start gives the semantics of English (in the actual world). Thus, it assumes that we should treat vagueness *within semantics*. But one might want to think of vagueness in a conceptually somewhat different way by just reinterpreting the partial and total valuation functions used in supervaluation theory: instead of saying that languages are vague, one can also say that it is our *use*, or *knowledge* of language that is imprecise. Lewis (1969), for instance, suggests that languages themselves are free of vagueness but that the linguistic conventions of a population, or the linguistic habits of a person, select not a point but a fuzzy region in the space of precise languages.¹⁰ A very similar view is taken in the epistemic approach of Williamson (1994): English, or its valuation function, is precise, but agents don't know exactly what this valuation function is. On such a view, borderline cases of *tall* are not individuals that are neither definitely tall nor definitely not tall *in English*, but rather individuals that some speakers of a language consistent with the linguistic convention of English consider to be tall, while others don't (on Lewis's meta-linguistic account), or individuals of which an agent doesn't know whether they count as being tall or not, although the agent knows their precise length (on the epistemic account). Notice that both analyses of vagueness are very much in line with supervaluation theory:¹¹ the partial valuation function from which the standard analysis starts still plays a role, if properly re-interpreted: it does not represent the semantics for English in the actual world, but rather what the *population* of speakers *agrees on*, or what an agent *knows about*, the actual interpretation function. Language users agree that the actual interpretation function is total, but disagree on, or are ignorant about, which one it is.

Notice that according to this re-interpretation, a total valuation can be thought of as a world that not only determines how the facts are (e.g. whether John's height is 1.80 meters or 1.70 meters), but also how a vague predicate like *tall* should be interpreted: whether somebody who is 1.80 meters should be considered to be tall or not. Worlds fulfill two roles, and those roles are exactly the roles a world can play according to Stalnaker's (1978) two-dimensional view on language discussed in section 2.4. If we fix the meanings of the expressions, a sentence expresses a (horizontal) proposition,

¹⁰Lewis (1970) takes this analysis of vagueness to be very similar to (what is now called) a supervaluation account. Burns (1991) argues (unconvincingly, we think) that the two are very different.

¹¹Neither proponents of the supervaluation account, nor proponents of the meta-linguistic or epistemic account would necessarily agree.

represented by a set of worlds, and if the actual world is a member of this set, what is said by the sentence is true, false otherwise. But if what is expressed by a (token of a) sentence depends on context, we can think of the world (together with the expression token) as determining how the expressions should be interpreted, and then it might be that in different worlds something different is said (i.e. different (horizontal) propositions are expressed) by the same sentential token. But if worlds fulfill the two roles suggested above, it seems natural to assume that they always fulfill the two roles *at the same time*. It follows that if we interpret a sentential token of a sentence ϕ in world w of which we consider it possible that it is the actual world, we use w both to determine *what is said* by ϕ , (denoted by $\llbracket\phi\rrbracket_w$), and to determining whether what is said by ϕ in w is *true* in w , i.e., whether $w \in \llbracket\phi\rrbracket_w$. The set of worlds denoted by $\{w \in W : w \in \llbracket\phi\rrbracket_w\}$ is called the *diagonal proposition* by Stalnaker (1978). The diagonal proposition expressed by a sentence is crucial to explain the so-called *evaluative meaning* of vague predicates. As noted by Barker (1992), if one says that ‘John is tall’, one can make two kinds of statements: a *descriptive* one saying that John is above the relevant cut-off point (if it is clear in a context what the cut-off point for being tall is) and a *metalinguistic* one saying that the cut-off point for being tall is below the length of John (if it is clear in a context what John’s height is). The latter involves the evaluative meaning of *tall* and can be accounted for straightforwardly in terms of diagonalization.

3 Meaning and Use

3.1 Speech acts

In the sections until now we have concentrated on declarative sentences, and assumed and defended a truth conditional analysis of sentence meaning. This is in accordance with the assumption that the only, or at least primary, aim of language is to represent and communicate factual information, information that can be true or false. Around the middle of the last century, however, this assumption was seriously challenged by philosophers like Wittgenstein and Austin. Both stressed that making factual statements is only one thing we do with language, and that we should study our use of language from a more general perspective of human action and behavior. The moment we leave the realm of declarative sentences, this seems an obvious move. By

using *imperative* and *interrogative* sentences like ‘Close the window!’ and ‘Is the window closed?’ we don’t describe an actual state of affairs, but rather make a *command* or ask a *question* in order to influence the behavior of our interlocutors. But, as pointed out by Austin, even for a whole class of declarative sentences the truth conditional analysis already seems unnatural. For instance, if John the judge says to the suspect ‘I sentence you to death’, to his brother ‘I bet that Ajax will win’, or to his wife ‘I promise to take care of our child’, his main purpose doesn’t seem to describe a state of affairs. What he seems to have done, rather, is to *change* the world: by his use of the sentence *a sentence*, *a bet*, or *a promise* came into existence that wasn’t really there before. Although Austin argued that for such so-called *performative* (uses of) sentences the question of truth or falsity doesn’t even arise, he noted that they can be used *feliculously* only if some *appropriateness conditions* are met. The utterance of ‘I sentence you to death’, for instance, only gives rise to a real sentence in case John makes it *as* a judge, and in a country where the death penalty is in use. But, of course, not only Austin’s performative sentences give rise to such appropriateness conditions for their successful use, imperative and interrogative sentences do so as well. The utterance of an imperative sentence like ‘Close the window!’, for instance, won’t be very successful in case the speaker has no authority over the hearer. Thus, the main point of all these sentences is to *change*, rather than to describe, the actual state of affairs, but this can be done successfully only if certain conditions are met. But as realized soon by Austin and others, this holds not only for imperative and interrogative sentences, together with the explicitly performative sentences of the type mentioned above, it is also true for standard declarative sentences like ‘The window is closed’. Although this sentence has truth conditions, a speaker uses it in an *assertion* only to make a point, and he can’t be successful in doing so when certain appropriateness conditions are not met. For instance, the speaker won’t make a point with his assertion in case it is already common knowledge among his conversational partners that the window is closed.

We have stated above that we make commands or ask questions in order to influence the behavior of our interlocutors. This seems to be true for making assertions as well, although perhaps more indirectly by influencing the hearer’s beliefs. So we might say that a command, question, or assertion is successful if and only if the hearer indeed performs the by the speaker intended effect. The intended effects for commands and questions would then be, most naturally, complying to the order and answering the question. But

this intended effect depends very much on context. For instance, it might well be that one intended effect of my particular assertion using the sentence ‘It is cold’ is that you not only believe that it is, in fact, cold, but also that you close the window. In an influential article, Grice (1957) tried to determine, or define, what the speaker means by a sentence in terms of this intended effect in a hearer by means of the recognition of this intention. It is clear, however, that complying to the order, or answering the question, does not follow automatically when the hearer recognizes the speaker’s intention. If I tell you to close the window, you can refuse to do so, and still, intuitively, understand what I meant.¹² The same is true for assertions: you can understand what I mean when I say ‘It is cold’ without closing the window, or even accepting that it is, in fact, cold. So, if we want to determine what the speaker means by a sentence in terms of the automatic effect in a hearer that follows from the recognition of this intention, we have to think of a specific kind of effect: that what Austin and Searle call the *illocutionary effect*. What could such an illocutionary effect be? Well, if I tell you to close the window and you don’t, I still have communicated something when you recognized my intention, namely that I make it public, between you and me, that I want you to close the window. Similarly for my assertion that it is cold: even if you don’t close the window, or believe what I say, if you recognized my intention I still have made it public between you and me that I want you to believe (and so make it common ground) that it is cold. As the examples illustrate, commands and assertions have different illocutionary effects, the one involves what I want you to do, the other what I want you to believe, or become common ground. For this reason they are called different *illocutionary acts*, or *speech acts*.

The traditional problem for speech act analyses was to find interesting types of speech acts, and to find necessary and sufficient conditions for the successful performance of the act. In these traditional analyses it was assumed that one could make a strict separation between what is expressed by a sentence, and the speech act performed by it. For instance, it was assumed that the sentences ‘Close the window!’, ‘Is the window closed?’, and ‘The window is closed’ all express the same proposition, namely that the window is closed, but that this proposition is used in different speech acts: a command, a question, and an assertion, respectively. Searle (1969) claims that assuming this hypothesis has many advantages. For instance, it allows us to

¹²Say something about that meaning of words is already presupposed too.

make a distinction between propositional and illocutionary negation. When a negation is applied to a proposition, it just results in another proposition but leaves the character of the illocutionary act unchanged. When a negation is applied to the illocutionary act, on the other hand, the proposition remains the same, but the illocutionary act changes. A negation of an assertion, for instance, gives rise to a *denial*, and the negation of a promise gives rise to the *refusal* to make a promise.

More recently, speech act theorists concentrated themselves on the *essential effects* of speech acts, and these effects are analyzed in terms of how the speech act changes the conversational situation. In the remainders of this section I will discuss assertions, questions, and commands and permissions, but will highlight within those discussions three separate issues. In the discussions of assertions I will take issue on whether we really *can* separate what is expressed by a sentence, its content, from its illocutionary force. Here I will also discuss some appropriateness conditions for making a successful assertion. When talking about questions I will discuss whether the content of an interrogative sentence is really as close to the content of an assertive sentence as traditional speech act analysis suggests, and what it means for the standardly assumed autonomy of semantics with respect to pragmatics. When looking at commands and permissions, finally, I will discuss whether permission sentences should best be treated as assertions or as imperatives in order to account for their well-known *free choice* performative effects.

3.2 Presupposition as a felicity condition

On the assumption that the primary aim of language is to represent factual information, all that counts for the interpretation of a sentence is its truth value (in a world). We have assumed above that the sentence ‘John came’ is true (in a world) if the referent of ‘John’ actually came, and false if this referent did not come. But what if the name ‘John’ has no referent? One natural reaction is to widen the concept of falsity: the sentence is true if the referent of the name actually came, and false otherwise. Unfortunately, as already observed by Frege (1898), this gives rise to the counterintuitive prediction that the negation of the sentence would not be ‘John did not come’, but rather ‘John did not come, or the name ‘John’ (as used by the speaker) has no reference’. Strawson (1950) famously proposed to solve the puzzle by claiming that if the referential expression has no reference, the sentence is neither true nor false. In order for ‘John came’, or ‘The king of

France is bald', to have a classical truth value (1 or 0), the referential terms occurring in it are required, or *presupposed*, to have a reference. In case of reference failure, the sentence in which the term occurs has a no classical truth value, though perhaps a non-classical one (i.e., *). After Strawson, linguists extended the notion of presupposition from referential terms to other types of expressions, including factual verbs like 'regret' and 'know', aspectual verbs like 'stop', and particles like 'even' and 'too'. Sentences like 'Mary *knows* that John came' and 'John came *too*', for instance, are said to presuppose that John came, and that somebody different from John came, respectively. These sentences would neither be true nor false, in case their presuppositions are not met.

So far, we have looked at simple sentences, but what about complex ones? If the truth value of a complex sentence involving a truth conditional connective is determined from the truth values of its parts, the most natural way of dealing with truth value gaps is to claim that any complex sentence inherits the true value gap of any of its parts. It is trivial to extend a two-valued logic into a three valued one which would have this result. But thinking of presupposition failure as having a truth-value gap gives rise to the prediction that whenever a simple sentence gives rise to a presupposition, any complex expression in which this simple sentence occurs gives rise to this presupposition as well. This prediction, however, is not in accordance with our intuitions: 'Mary came and John came *too*' doesn't give rise to a presupposition, and, indeed, the sentence seems false in case nobody different from John came. One way to solve this problem is to come up with a new three-valued logic, where conjunction, for instance, does not give rise to a symmetric truth table, but to an *asymmetric* one instead: 'A and B' can be said to be false, rather than neither true nor false, when 'A' is false and 'B' neither true nor false.¹³ Although such an account would be more in accordance with our intuitions, this way of solving the problem seems to be rather *ad hoc*. Furthermore, it seems rather dubious whether we should account for presuppositions solely in terms of a third truth value. For one thing, because some people have argued that some sentences (like '*Even* John came', or 'Mary does not *regret* that John came') can be true although their presuppositions are not met. For another, because it is rather doubtful whether we have firm, and theory neutral, intuitions that tell us whether a sentence is neither true nor false in a particular circumstance at all.

¹³This is Peters' (1977) truth table of conjunction.

Both of these problems can be met when we think of language, and of presupposition, from a more general perspective. Once we assume that the primary aim of using declarative sentences is to communicate, rather than just to represent, factual information in order to influence the beliefs or actions of ones' conversational partners, we can think of presupposition as just a special kind of felicity, or appropriateness, condition for the successful use of a speech act. In Stalnaker's (1978) classical analysis, an assertion of a declarative sentence ' ϕ ' is successful just in case it increments, or updates, what is presumed to be commonly believed with the content of the assertion. Thus, an assertion of ' ϕ ' is made with respect to the context of what is taken to be commonly believed, represented by K , and its pragmatic effect is that this context is updated from K to $Upd(\phi, K)$. If we assume that the context can be represented by a set of possible worlds, this means that $Upd(\phi, K) = K \cap [\phi]$, where $[\phi]$ is the proposition denoted by ' ϕ '. If the aim of an assertion is to update the context, this context has to satisfy certain conditions in order for the assertion to be successful. For instance, the context should not yet entail the proposition expressed by ' ϕ ', because then the assertion would not change the context, and thus have no pragmatic effect. But a further constraint now follows naturally. Notice that on a common sense of 'presupposition', what is presupposed by a speaker is just what he takes to be common ground between the participants of a conversation. On this view, it is primarily speakers that presuppose something. However, a sentence might presuppose something as well: we can say that sentence ' ϕ ' presupposes P just in case ' ϕ ' can be appropriately uttered by a speaker only if he presumes it to be common ground that P is the case. But this means that the assertion of ' ϕ ' puts a constraint on the contexts in which it can be used appropriately: it has to be a context K that already entails, or satisfies, P .

This speech act analysis of presuppositions can, arguably, solve the problems discussed above for a truth-conditional analysis of presuppositions (cf. Stalnaker, 1974). First, it can account for the fact that a complex sentence like 'Mary came and John came *too*' doesn't give rise to a presupposition, although the second conjunct does. The reason is that it seems reasonable to assume that the context of interpretation of the second conjunct is not the initial context, K , but rather the initial context updated with the content of the first conjunct. Because in this updated context the presupposition of the second conjunct is satisfied, even if this is not the case in the initial context, the utterance of the conjunctive sentence doesn't put any constraint

on initial contexts, and thus doesn't give rise to a presupposition. Second, the speech act analysis of presuppositions can, in principle, account for the intuition that a sentence can be true (or false) in the actual world, although in this world the presupposition is not met. The reason is that the speaker might presuppose something that is actually false, and so the actual world is not an element of the context.

But is the above picture not simply mistaken? Isn't it obvious that a sentence can be used that, intuitively, gives rise to a presupposition, although the speaker does not take it to be common ground that this 'linguistic' presupposition is true before he made the assertion? Indeed, Mary can say that she *regrets* that John did not come to convey the new information that John did, in fact, not come. But perhaps not all information that is conveyed by a sentence should be accounted for by semantics. Perhaps what is called presuppositional inference is one such case. And this makes sense from the pragmatic point of view: In case it is commonly known that a sentence ϕ can *normally* be asserted appropriately only in case certain information ψ is already taken for granted by the participants of the conversation, it becomes possible to *exploit* this knowledge by using ϕ to *pretend* that ψ is already assumed. A speaker can pretend to take something to be already common ground and thereby, indirectly, convey new information. For this to be possible, however, it is required that this pretense is the exception, rather than the rule. So we see that our appealing pragmatic picture of presuppositions can be appropriate to the extent that in most, or at least in typical, conversational situations in which the speaker uses a sentence that gives rise to a presupposition, this presupposition is already common ground.

3.3 Questions and the autonomy of semantics with respect to pragmatics

Just like an *assertion*, also a *question* is a speech act: it is something we *do* with a sentence. However, the sentences we typically use to ask a question differ from the sentences we typically use to make an assertion. While assertions correspond with *declarative* sentences, questions correspond with *interrogative* sentences. But this correspondance is not complete: declarative sentences, for instance, can be used not only to make assertions, but also to ask questions. This is typically the case for declarative sentences with rising intonation. Suppose that we can determine whether a sentence is used as a

question or not. Then we need to know what is the meaning of the sentence, and what is the pragmatic effect.

So, what is the meaning of a question? It seems natural that a question like ‘Did John come?’ also involves a proposition, but how should this involvement be spelled out? In traditional speech act theory it was assumed that the question simply expresses the proposition that John came, i.e., that the *meaning* of a question is just a proposition, but that this proposition is just used differently than in an assertion. So how is this proposition used in a question, i.e., what is the pragmatic effect of ‘Did John come?’ What the speaker expresses with the sentence is that he wants to know what is the correct and satisfying true answer. But what is a satisfying answer to a question? For yes-no questions this seems obvious: ‘yes’ and ‘no’, or just the proposition expressed by a yes-no question and its negation. So what the speaker then wants to know is whether, of all possible worlds that he takes to be live options, the actual world is one where the proposition expressed by the question is true, or not. But this means that the essential effect of a yes-no question is to introduce to the context the issue of whether the proposition expressed by the sentence is true. Thus, what the question does is to divide the worlds of the context into those where the proposition holds and those where it does not, i.e., it *partitions* the context.

Let us say, following traditional speech act theory, that the meaning of the question ‘Did John come?’ is the set of possible worlds in which John came. Equivalently, this is just the following function from worlds to the value ‘true’ if John came in that world, and the value ‘false’ otherwise: $\lambda w[\text{John came in } w]$. The pragmatic effect of the question with respect to context K , $Upd(\text{Did John come}, K)$, can now simply be modeled as the following partition of K .

$$\{\{v \in K : \lambda w[\text{John came in } w](v) = \lambda w[\text{John came in } w](u)\} \mid u \in K\},$$

Because of the correspondence between partitions and equivalence relations, the pragmatic meaning of ‘Did John come?’ can also be given by the equivalence relation $\lambda v \lambda w[\text{John came in } w \text{ iff John came in } v]$ on K .

To account for *wh*-questions along the same lines, it seems we need propositional functions and not complete propositions, because in contrast to yes-no questions a speaker asking a *wh*-question does not express a complete proposition. So let us assume that the meaning of a question like ‘Who came?’ is the function, $\lambda w \lambda x[x \text{ came in } w]$, that when applied to a world

and an individual assigns the value ‘true’ if the individual came in that world, and the value ‘false’ otherwise. To determine the pragmatic effect in a similar way as for yes-no questions, we have to know what counts as a satisfying answer to the question. In principle, many answers can be given to the question, but some are more natural than others. But we don’t care about natural answers here, but only about the kind of answers that would fully satisfy the questioner. A fully satisfying answer seems to be one where he learns the *complete* answer, and knows afterwards who came. Groenendijk & Stokhof (1982) have argued that to know who came, the agent needs to know of *each* single individual *whether* he or she came. On this proposal, the pragmatic effect of a question is to give the set of all possible complete answers. Notice that such complete answers exclude each other, and, given that for each world there is one complete answer true in it, the pragmatic effect of ‘Who came?’ with respect to context K , $Upd(\text{Who came?}, K)$, thus gives rise to a set of propositions which partitions context K as well, or, equivalently, to the equivalence relation on K below.

$$\{\{v \in K \mid \lambda w \lambda x [x \text{ came in } w](v) = \lambda w \lambda x [x \text{ came in } w](u)\} \mid u \in K\}$$

$$\lambda v \lambda w [\lambda x [x \text{ came in } w] = \lambda x [x \text{ came in } v]]$$

Notice that on this analysis, the pragmatic effects of questions all give rise to partitions. This allows us to define an entailment relation between questions, in terms of their pragmatic meanings, or effects. Suppose that Q^K and Q'^K are the pragmatic effects of two interrogative sentences with respect to context K . Then we might say that the first interrogative sentence *pragmatically entails* the second interrogative sentence just in case for every proposition that is an element of the partition Q^K there is a proposition that is an element of the partition Q'^K such that the former is a subset of the latter. More formally, Q^K pragmatically entails Q'^K iff $\forall q \in Q^K : \exists q' \in Q'^K : q \subseteq q'$.¹⁴ This abstract characterization seems to make sense as well. For instance, it predicts that the question ‘Did John come?’ is pragmatically entailed by ‘Who came?’, because any (complete) answer to the latter question also completely answers the former question. This prediction seems to be in accordance with our intuitions.

So it seems that the traditional speech act analysis of questions is quite appealing. Still, semanticists typically don’t adopt this analysis. The rea-

¹⁴One might also abstract away from the context, or from the model, but I will leave that to the reader.

son for this is that if we assume that the meaning of a sentence should be determined in terms of the meanings of their parts, the analysis can't account for *embedded* sentences like 'Mary knows who came'. On the most natural analysis of sentences like 'Mary knows that John came', the verb 'know' denotes a relation between an agent and the proposition expressed by the embedded sentence 'John came'. Unfortunately, on the present analysis, the semantic meaning of 'who came' is a propositional function, rather than a proposition. So, either we have to assume that the meaning of the verb *know* is ambiguous between the standard one that involves a proposition, and a question-meaning that involves a propositional function, or we have to give up the assumption that the meaning of 'Mary knows who came' can be compositionally determined in terms of the meanings of its parts. In fact, a simple ambiguity between a propositional and a question-meaning of 'know' is not going to be enough, for on the standard speech act analysis the meaning of a question cannot only be a propositional function, but also a propositional relation, or a proposition, as in 'Mary knows whether John came'. But even if we assume that the meaning of the verb 'know' is multiple ambiguous, it is still not clear how it could account for the intuitively correct truth conditions of sentences like 'Mary knows who came' and 'Mary knows whether John came'. According to almost everybody's intuition, the latter sentence is true if and only if Mary knows *that* John came if John in fact came, and Mary knows *that* John did not come if John did in fact not come. But how can we derive these truth conditions if we assume that the meaning of the embedded question is just the proposition *that* John came? Similarly for embedded *wh*-questions: if the semantic meaning of such a question is just a propositional function, how can we account for the intuition we have discussed above that 'Mary knows who came' is true if and only if Mary knows for each individual whether that individual came?

Fortunately, there seems to be a very straightforward solution to this problem. Just assume that to account for the truth conditions that involve embedded questions we shouldn't look at the *semantic* meanings of these embedded questions, but rather at their *pragmatic meanings*. Remember that according to the traditional speech act analysis one might think of the pragmatic meaning of a question as a partition on, or an equivalence relation between, possible worlds. If 'whether John came' and 'who came' have such equivalence relations as their pragmatic meanings, these relations between worlds give rise to propositions if they are applied to a world. Now assume that to determine the truth conditions, to check whether 'Mary knows Q' is

true in a world w – where Q is the embedded question –, we just look at the proposition determined by the pragmatic meaning of Q applied to w . Now we can assume that the meaning of ‘know’ is just a relation between individuals and propositions, and we correctly predict that ‘Mary knows whether John came’ is true just in case Mary knows *that* John came if John in fact came, and Mary knows *that* John did not come if John did in fact not come. One can check that on this procedure also ‘Mary knows who came’ receives the intuitively correct truth conditions.

There is, in fact, not so much to say against this analysis, except for the following methodological complaint: once we take embedded questions into account, we cannot determine the *semantic* meaning of the whole sentence solely in terms of the *semantic* meanings of their parts and the way they are formed together. Instead, we have to take the *pragmatic* meaning into account as well. This, obviously, contradicts Searle’s explicitly mentioned assumption that we can determine the proposition expressed by a sentence without the mentioning of illocutionary force, and more in general it would lead us to give up the assumption that semantics is *autonomous* with respect to pragmatics.

Suppose that we want to keep semantics autonomous with respect to pragmatics, how should we proceed? Given the above suggested analysis, this is easy to see, and it gives rise to Groenendijk & Stokhof’s (1982) analysis.¹⁵ Just as it is normally assumed that you know the meaning of a declarative sentence when you know under which circumstances this sentence is true, Hamblin (1958) argues that you know the meaning of a question when you know what counts as a satisfying answer to the question. If we say, as we did above, that only complete answers count as satisfying answers, this means that the semantic meaning of a question on this analysis is just the same as the pragmatic meaning of the same question on the suggested analysis above. It follows that the meaning of a question is a partition, or an equivalence relation, and that we can account for our intuitions concerning entailments between questions in terms of their semantic, rather than their pragmatic, meanings.

But how should we now account for the *pragmatic* effect of questions, and

¹⁵In fact, the above sketched pragmatic analysis was, of course, modeled on Groenendijk & Stokhof’s semantic analysis. To be sure, there are other semantic analyses of questions, and they give rise to different – and some would say better – empirical predictions. But my main concern here is methodological rather than empirical, and so I won’t discuss those alternative semantic analyses here.

what do we do with *embedded* questions? As for the pragmatic effect things are easy: we can just say that the effect of a question when used in a context is that it (further) partitions this context. As for embedded questions, we now say that they have the same semantic meanings as their unembedded counterparts, and that *Mary knows whether John came*, for instance, is true if and only if Mary knows the proposition denoted by the semantic meaning of the embedded question applied to the actual world. This obviously gives rise to the same, and thus intuitively correct, semantic meaning of the whole sentence as on the analysis described above, except that now we don't have to give up the assumption that semantics is autonomous with respect to pragmatics.

3.4 Permissions and the free choice effect

3.4.1 The problem of free choice permissions

According to Austin's classical analysis of speech acts, sentences of the form *You must/may do ϕ* are not used to describe a states of affairs. In terms of the language game between master and slave as described by Lewis (1970/9), they are typically used by one person, the master, to command or permit another person, the slave, to do certain things.

How should we account for these so-called *performative* effects of the sentences used by the master? One proposal might be to say that command and permission sentences are *assertorically* used, but that the performative effect is accounted for in an *indirect* way, due to the fact that we *learn*, or realize, more about the world. One of the things one might learn about the world is what is demanded and permitted. A truth conditional analysis of what is demanded and permitted is given in deontic logic. Standard deontic logic (SDL) was based on the same principles as classical modal logic.¹⁶ Where normal modal logic has the operators \Box and \Diamond standing for necessity and possibility, SDL has the two operators O and P , standing for *ought* or *obliged* and for *permission*, respectively. Model theoretically, we say that $O(\phi)$ is true in w iff in all ideal worlds accessible from w , ϕ is true, and that its dual $P(\phi)$ is true iff ϕ is consistent with this set of all ideal worlds, i.e. if there is at least one ideal world accessible from w in which ϕ is true. The set of ideal worlds in w will be denoted by $P(w)$, and is known as

¹⁶There are other truth conditional analyses of deontic concepts, of course, but we won't go into that here.

the *permissibility set*. We might now propose that the performative effect of command and permission sentences is due to the fact that only after a command or permission sentence is used by the master, the slave *knows* that he is obliged/permitted to do something, by having eliminated worlds with inappropriate permission sets, and acts accordingly.

This assertoric analysis seems appropriate for some uses of command and permission sentences, but it has always taken to be problematic whether the performative effect of all permission sentences should be accounted for in the *epistemic* way sketched above. Consider the sentence ‘You may take the apple or take the pear’. According to standard deontic logic, this sentence follows from both ‘You may take the apple’ and from ‘You may take the pear’, and neither of them follows from the first disjunctive permission. In a sense this is how things should be because, as observed by Kamp (1979), there is nothing problematic with the assertion of ‘You may take the apple or the pear, but I don’t know which.’ On the other hand, however, we can intuitively infer both ‘You may take the apple’ and ‘You may take the pear’ from the disjunctive permission sentence. How could we possibly account for this latter free choice inference if the latter also can be inferred from the former two?¹⁷ In the following I will discuss two proposed solutions to this problem: (i) a performative analysis, and (ii) an analysis that explains the free choice inference as a conversational implicature.

3.4.2 The performative analysis of imperatives

The natural alternative to the assertoric analysis of obligation and permission sentences is the *performative* one involving a master and his slave. According to the performative analyses of Lewis (1970/9) and Kamp (1973), command and permission sentences are not primarily used to make true assertions about the world, but rather they are made by the master to *change* that what the slave is obliged/permitted to do.¹⁸ With some feeling for Amster-

¹⁷Many authors have discussed this puzzle, and this typically involves dropping the standard truth conditional analysis of ‘or’, or the standard analysis of modals such that the disjunctive permission doesn’t follow anymore from any of ‘You may take the apple’ and ‘You may take the pear’. I won’t go into those more desperate attempts to solve the problem, and will only discuss (in my eyes) more appealing proposals that stay rather classical.

¹⁸Although Lewis (1970/9) and Kamp (1973) account for the effect of permission sentences in rather different ways, both might be called performative analyses in the sense that their effect is to change the permissibility set.

dam rhetorics, we might say that according to the performative analysis, we know the meaning of an imperative sentence, when we know how imperatives change permissibility sets.

According to this Lewis/Kamp account, if the master commands John to do ϕ by saying ‘You must do ϕ ’, or allows John to do ϕ by saying ‘You may do ϕ ’, it is typically not yet the case that the proposition expressed by ϕ is respectively a superset of, or consistent with, John’s permissibility set, P .¹⁹ However, the performative effect of the command/permission will be such that in the new context what is commanded is a superset of, and what is permitted is consistent with, the new permissibility set. Thus, in case the command or permission is not used vacuously, the permissibility set, P' , of the new context will be different from P , so that the obligation/permission sentence will be satisfied.

But if knowing the meaning of an imperative means that you have to know how the imperative changes the permissibility set, our problem is to say how command and permission sentences govern the change from the prior permissibility set, P , to the posterior one, P' .

For commands this problem seems to have an easy solution. If the command ‘You must do ϕ ’ is given by the master, the new, or posterior, set of permissible futures for John, P' , is simply $P \cap [\phi]$, where $[\phi]$ denotes the proposition expressed by ϕ .²⁰ However, things are more complicated for permission sentences. It is clear that if ϕ is allowed, P' should be a superset of P such that $P' \cap [\phi] \neq \emptyset$. It is not clear, however, which ϕ -worlds should be added to P . Obviously, we cannot simply say that $P' = P \cup [\phi]$. By that suggestion, giving permission to ϕ would allow everything compatible with ϕ , which is certainly not what we want. But how then should the change from P to P' be determined if a permission is given? This is Lewis’s problem about permissions.

One possible way to solve Lewis’s problem about permissions is to assume that we not only have a set of best, or ideal, worlds, but also an ordering that says which non-ideal worlds are better than others. Thus, to account for the performative effects of commands and permissions, we need not only a set of ideal worlds, but rather a whole preference, or reprehensibility, ordering, \leq , on the set of all possible worlds. On the interpretation that $u \leq v$ iff v is at

¹⁹From now on I will assume in most of this paper that there is only one (global) permissibility set around.

²⁰What if the new command is incompatible with one or more of the earlier ones? In that case we might make use of change by revision to be discussed below.

least as reprehensible as u , it is natural to assume that this relation should be reflexive, transitive, and connected.²¹ In terms of this preference order on possible worlds we can determine the ideal set P as the set of minimal elements of the relation \leq :

$$P \stackrel{\text{def}}{=} \{v \in W \mid \forall u : v \leq u\}$$

In terms of this set of ideal worlds we can, as before, determine of course whether according to the present state ϕ is obligatory or just permitted. For instance, ϕ is obligatory iff $P \subseteq [\phi]$.

But this ordering relation contains more information than just what the set P of ideal worlds is, and in terms of this extra information we can determine the new permissibility set P' . If the master permits the slave to make ϕ true, we can assume that P contains no ϕ -worlds, i.e. none of the ϕ -worlds is ideal. But some ϕ -worlds are still better than other ϕ -worlds. We can now propose that the effect of allowing ϕ is that the best ϕ -worlds are added to the old permissibility set to figure as the new permissibility set. The best ϕ -worlds are the worlds ‘closest’ to the ‘ideal’ worlds P where ϕ is true. This set will be denoted as P_ϕ^* and defined in terms of the relation \leq as follows:

$$P_\phi^* \stackrel{\text{def}}{=} \{u \in [\phi] \mid \forall v \in [\phi] : u \leq v\}$$

To implement this suggestion, we can say that the change induced by the permission *You may do ϕ* is that the new permission set, P' , is just $P \cup P_\phi^*$.²² Thus, according to this proposal, command and permission sentences change a context of interpretation as follows (where I assume that John is the relevant agent, and P his permission state):

$$\begin{aligned} \text{Upd}(\text{Must}(\text{John}, \phi), P) &= P \cap [\phi]^{23} \\ \text{Upd}(\text{May}(\text{John}, \phi), P) &= P \cup P_\phi^* \end{aligned}$$

Note that according to our performative account it does not follow that for a permission sentence of the form ‘You may do ϕ or ψ ’ the slave can

²¹A relation R is *reflexive* if for all $w : R(w, w)$, it is *transitive* if for all w, v and u : if $R(w, v)$ and $R(v, u)$, then $R(w, u)$, and it is *connected* if for all w and v , $R(w, v)$ or $R(v, w)$.

²²This analysis of permission sentences was assumed by Kamp (1979) in his discussion of the performative analysis of permissions.

²³This is, in fact, just P_ϕ^* , where it is assumed that ϕ is compatible with P .

infer that according to the new permissibility set he is allowed to do any of the disjuncts. Still, the performative analysis can give an explanation why disjuncts are normally interpreted in this ‘free-choice’ way. To explain this, let me first define a deontic preference relation between propositions in terms of our reprehensibility relation between worlds, \leq . We can say that although both ϕ and ψ are incompatible with the set of ideal worlds, ϕ is still preferred to ψ , $\phi \preceq \psi$, iff the best ϕ -worlds are at least as close to the ideal worlds than the best ψ -worlds, $\exists v \in [\phi]$ and $\forall u \in [\psi] : v \leq u$. Then we can say that with respect to \leq , ϕ and ψ are equally reprehensible, $\phi \approx \psi$, iff $\phi \preceq \psi$ and $\psi \preceq \phi$. Because, as it turns out, it will be the case that $P_{\phi \vee \psi}^* = P_{\phi}^* \cup P_{\psi}^*$ iff $\phi \approx \psi$, we can now explain why normally disjunction elimination is allowed for permission sentences.²⁴ For simple disjunctive permission sentences like ‘You may do ϕ or ψ ’, it is not unreasonable to assume that when performatively used, the master has no strict preference for the one above the other. If we make the same assumption for command sentences, it follows that from ‘You may/must take the apple or the pear’ we can conclude that the speaker may take the apple and that he may take the pear.

This performative analysis gives rise to a number of problems, but for reasons of space I will consider only two. First, the analysis might be natural for permissions that are, intuitively, used performatively, but still can’t account for the intuition that the ‘free-choice’ effect results even if the sentence is used assertively. So, even if the analysis is correct, one still needs an analysis that makes the same predictions for assertively used permissions. But, then, perhaps, such an assertive analysis is all that one needs. The second problem is very similar: it seems that other sentences involving disjunction have ‘conjunctive’ readings as well, although for these examples it is neither performativity nor modality that seems to be crucial.²⁵ from ‘Several of my cousins had cherries or strawberries’ we naturally infer that some of the cousins had cherries and some had strawberries. These problems suggest that we (at least also) need a general assertive analysis of free choice inferences. In the following section I suggest to account for the inferences as conversational implicatures.

²⁴It is well possible to give a performative analysis of permission sentences where the free choice effect comes about without requiring that both disjuncts are equally reprehensible, but I won’t go into that here.

²⁵I learned these examples from Regine Eckhardt.

3.5 A pragmatic analysis of Free Choice

The fact that there is nothing wrong with the assertion of ‘You may take the apple or the pear, but I don’t know which’ suggests that the free choice permission inference is cancellable, and should be accounted for as a Gricean conversational implicature.

3.5.1 Gricean implicatures

Traditionally, the semantic meaning of natural language expressions like ‘and’, ‘or’, ‘every’, ‘some’, ‘believe’, and ‘possibly’ has been analyzed in terms of their intuitive analogs in classical logic: ‘ \wedge ’, ‘ \vee ’, ‘ \forall ’, ‘ \exists ’, ‘ \square ’, and ‘ \diamond ’, respectively. However, in many contexts these expressions receive interpretations that are different from what is predicted by this approach to their semantics. In most circumstances, for instance, we infer from the assertion of ‘John is came *or* Mary came’ that John and Mary didn’t come together, and from ‘It is *possible* that John came’ that it is *not necessary* that John came.

How should these inferences be accounted for? Grice (1967) argued that the above inferences should not be accounted for within a semantic analysis, but should be accounted for in terms of general principles of rational communication. Grice assumes a theoretical distinction within the ‘total significance’ of a linguistic utterance between what the speaker *explicitly said* and what he has merely *implicated*. What has been said is supposed to be based purely on the *conventional* meaning of a sentence, and is the subject of compositional semantics. What is implicitly conveyed belongs to the realm of pragmatics and depends also on facts about the utterance situation, the linguistic context, and the goals and preferences of the interlocutors of the conversation. What is implicitly conveyed, or *conversationally implicated* can be determined, or so is proposed, on the basis of Grice’s *cooperative principle*: the assumption that speakers are maximally efficient rational cooperative language users. Grice comes up with a list of four rules of thumb – the maxims of *quality*, *quantity*, *relevance*, and *manner* – that specify what participants have to do in order to satisfy this principle. They should speak sincerely, relevantly, and clearly, and should provide sufficient information.

Over the years many phenomena have been explained in terms of the Gricean maxims of conversation. Horn (1972) and especially Gazdar (1979) proposed to formalize Grice’s suggestions in order to turn informal pragmat-

ics into a predictive theory. They concentrated on Grice’s maxim of quality and his first submaxim of quantity. Grice’s maxim of quality says, roughly speaking, that the speaker always *knows* (or believes) what he says, while his first submaxim of quantity (and Relevance) assumes that the speaker makes his contribution as informative as required. Obviously, to implement these maxims, we need to take the knowledge state of speakers into account.

Formalizing that the speaker obeys quality is not that difficult: If our designated speaker utters ϕ , we simply assume that the speaker’s knowledge state entails ϕ , and thus that $\mathbf{K}\phi$ is true. Thinking of \mathcal{S} as the set of knowledge states, quality demands that a speaker of ‘ ϕ ’ is in one of the following knowledge state: $\{s \in \mathcal{S} | s \subseteq [\phi]\}$. To account for the first subclause of the maxim of quantity that demands speakers to convey all (relevant) information they possess, we are going to select among those states where the speaker knows her utterance to be true the states where she has least additional relevant knowledge. This is formalized by defining an order on epistemic states and then select minimal elements of this order. The order compares the relevant knowledge of the speaker and we select minimal elements in the set $\{s \in \mathcal{S} | s \subseteq [\phi]\}$. How much relevant knowledge a speaker has is taken to be represented by how many of a class of alternative sentences she knows to hold. Let us assume that if $\phi = \text{‘[John]}_{\mathcal{F}} \text{ smokes’}$, for instance, the set of alternatives contains sentences like ‘John smokes’, ‘Mary smokes’ and ‘Bill smokes’ as well as the conjunctive and disjunctive combinations of them. Now we say that the speaker has less relevant knowledge in state s than in s' , $s <_{\mathbf{K}Alt(\phi)}^{\mathbf{K}} s'$, iff the set of alternative sentences known in the former state is a proper subset of the set of alternative sentences known in the latter state:

Definition 1 (*Ordering knowledge states*)

$$s \leq_{\mathbf{K}Alt(\phi)}^{\mathbf{K}} s' \quad \text{iff} \quad \{\psi \in Alt(\phi) : s \subseteq [\psi]\} \subseteq \{\psi \in Alt(\phi) : s' \subseteq [\psi]\}.$$

Now we define the Gricean interpretation of ϕ as the set of minimal models where the speaker knows ϕ with respect to the set of alternatives $Alt(\phi)$.

Definition 2 (*A Gricean Interpretation*)

$$[Grice]^{\mathcal{S}}(\phi, Alt(\phi)) = \{s \subseteq [\phi]^{\mathcal{S}} : \forall s' \subseteq [\phi]^{\mathcal{S}} : s \leq_{\mathbf{K}Alt(\phi)}^{\mathbf{K}} s'\}.$$

According to this interpretation function, if the speaker utters ‘[John] $_{\mathcal{F}}$ came’ we conclude that the speaker knows that John came, but not that Mary came,

and if she utters ‘[John or Mary]_F came’ we conclude that the speaker does not know of anybody that he or she came. This is a nice result, but in many cases we conclude something stronger: in the first example that Mary, Bill, and all the other relevant individuals *did not* come, and the same for the second example, except that now this is not true anymore for Mary. How do we account for this extra inference in terms of our richer modal-logical setting?

In van Rooij & Schulz (2004) it is shown that this can be accounted for by assuming that speakers, in addition to obeying the Gricean maxims, are *maximally competent* (as far as this is consistent with obeying these maxims).²⁶ This can be described by selecting among the elements of $[Grice]^S(\phi, Alt(\phi))$, the ones where the competence of the speaker is maximal. To account for this we need a new order that compares the competence of the speaker. This order is described in definition 3.

Definition 3 (*Ordering by consistency statements*)

$$s <_{Alt(\phi)}^P s' \text{ iff } \{\psi \in Alt(\phi) : s \cap [\psi] \neq \emptyset\} \subset \{\psi \in Alt(\phi) : s' \cap [\psi] \neq \emptyset\}.$$

The minimal models in this ordering are those states where the speaker knows *most* about the alternatives. Now, finally, we define the function $[Comp]^S(X, Alt(\phi))$ (*Comp* stands for competence) by selecting the minimal elements in X according to the ordering $<_{Alt(\phi)}^P$:

Definition 4 (*Maximizing competence*)

$$[Comp]^S(X, Alt(\phi)) = \{s \in X : \neg \exists s' \in X : s' <_{Alt(\phi)}^P s\}.$$

If we now apply $[Comp]^S$ to $[Grice]^S(\phi, Alt(\phi))$, where ϕ is a sentence like ‘[John]_F came’ or ‘[John]_F came or [Mary]_F came’, we see that from the first we can conclude that the speaker knows that Mary and Sue did not come, while from the second that the speaker knows that Sue did not come, but also that it is not the case that John and Mare came together. So, in this way we have accounted for the exclusive reading of ‘John came *or* Mary came’.

One can show (cf. Van Rooij & Schulz, 2004, Spector, 2006) that if we apply $[Comp]^S$ to $[Grice]^S(\phi, Alt(\phi))$ we derive exactly the same implicatures as

²⁶The same idea can be found also in Spector (2006).

we can derive using exhaustive interpretation. The exhaustive interpretation of ϕ with respect to its alternatives $Alt(\phi)$ is defined as follows:

$$exh(\phi, Alt(\phi)) = \{w \in [\phi]^W : \neg \exists v \in [\phi]^W : v <_{Alt(\phi)} w\},$$

where $v <_{Alt(\phi)} w$ iff $\forall \psi \in Alt(\phi) : v \in [\psi]^W \rightarrow w \in [\psi]^W$.

How does our Gricean analysis account for the inference that it is not necessary that John came, if the speaker asserted ‘It is *possible* that John came’? This is quite straightforward. We can just assume that one of the alternatives of a sentence of the form ‘ $\diamond\phi$ ’ is the sentence ‘ $\Box\phi$ ’. It is easy to see that from $[Grice]^S(\diamond\phi, Alt(\diamond\phi))$ we can conclude that the speaker doesn’t know that $\Box\phi$ is true. If we then assume that, in addition, the speaker is maximally competent on the alternatives, it follows that the speaker knows that $\Box\phi$ is false, and thus that ϕ is, in fact, not necessary the case.

3.5.2 A pragmatic analysis of free choice permission

Consider again sentences of the form ‘You/John may take the apple or the pear’, both represented by $\diamond(\phi \vee \psi)$. Let us assume, for simplicity, that the alternatives of the embedded clause are just ϕ and ψ itself, together with their conjunction and disjunction. Let us also assume that the set of alternatives to a sentence of the form ‘ $\diamond\phi$ ’ is just given by the set $\{\diamond\psi : \psi \in Alt(\phi)\} \cup \{\Box\psi : \psi \in Alt(\phi)\}$. If we now apply the Gricean interpretation rule of the previous section, it is easy to see that things go wrong: we see that for each alternative to $\diamond(\phi \vee \psi)$ (except the sentence itself, of course), there is a model that makes $\diamond(\phi \vee \psi)$ true but not this alternative, but that there is no model that falsifies all these alternatives together: neither $\diamond\phi$ nor $\diamond\psi$ has to be true in order for $\diamond(\phi \vee \psi)$ to be true, but they cannot be false both. There are various ways to go to solve this problem: either change what counts as an alternative, or the Gricean interpretation rule *Grice*.

According to the first alternative, proposed by Schulz (2003), the set of alternatives to a sentence of the form ‘ $\diamond\phi$ ’ is just given by the set $\{\Box\psi : \psi \in Alt(\phi)\} \cup \{\Box\neg\psi : \psi \in Alt(\phi)\}$. First, notice that by applying *Grice* to a sentence of the form ‘ $\diamond(\phi \vee \psi)$ ’ it immediately follows that the speaker knows neither $\Box\neg\phi$ nor $\Box\neg\psi$, in formulas, $\neg\mathbf{K}\Box\neg\phi$ and $\neg\mathbf{K}\Box\neg\psi$. What we would like is that from here we derive the free choice reading: $\diamond\phi$ and $\diamond\psi$, which would follow from $\mathbf{K}\neg\Box\neg\phi$ and $\mathbf{K}\neg\Box\neg\psi$. Of course, this doesn’t follow yet, because it might be that the speaker does not know what the agent

may or must do.²⁷ But now assume that the speaker is competent on this.²⁸ Intuitively, this assumption means that the speaker thinks it is *possible* that the agent can or must do a if and only if the speaker *knows* that the agent can or must do a . In formulas: $\mathbf{P}\Box\phi \equiv \mathbf{K}\Box\phi$ and $\mathbf{P}\Diamond\phi \equiv \mathbf{K}\Diamond\phi$. This assumption is completely natural for performatively used permission sentences, because in that case the speaker is the authority on what is permitted. But for assertively used sentences it is sometimes natural to make this assumption as well. Remember that after applying *Grice*, the minimal models falsify $\mathbf{K}\Box\neg\phi$ and $\mathbf{K}\Box\neg\psi$, which means that $\mathbf{P}\neg\Box\neg\phi$ and $\mathbf{P}\neg\Box\neg\psi$ have to be true. The latter, in turn, are equivalent to $\mathbf{P}\Diamond\phi$ and $\mathbf{P}\Diamond\psi$. By competence we can now immediately conclude to $\mathbf{K}\Diamond\phi$ and $\mathbf{K}\Diamond\psi$, from which we can derive $\Diamond\phi$ and $\Diamond\psi$, because knowledge implies truth. Thus, following Schulz' (2003) minimal modal analysis, we get the free choice effect as a pragmatic inference.²⁹

One natural step would be to say that the speaker is competent on who John thinks might have passed the examination. In that case, the above minimal state disappears, and we will end up with *two* minimal states according to the $\leq_{\mathcal{L}}^{\mathbf{K}}$ ordering: One where the speaker knows that only $\Diamond_j P(a)$

²⁷Notice, though, that this inference *does* follow if '□' and '◇' stand for epistemic *must* and epistemic *might*. This is so, because for the epistemic case we can safely assume that the speaker knows what he believes, which can be modeled by taking the epistemic accessibility relation to be fully introspective. This predicts correctly, because from 'Katrin might be at home or at work', it intuitively follows that, according to the speaker, Katrin might be at home, and that she might be at work.

²⁸Formally this is done by making a constraint on models: consider only models where the speaker knows the (deontic) accessibility relation of the agent.

²⁹It is easy to see that this analysis can account for the 'free choice'-inference of the existential sentence as well: that from 'Several of my cousins had cherries or strawberries' we naturally infer that some of the cousins had cherries and some had strawberries. First we assume that the sentence is represented by something like $\exists x[Px \wedge (Qx \vee Rx)]$. Then we take the alternatives of this existential formula to be universal formula: $\forall x[Px \rightarrow Qx]$, $\forall x[Px \rightarrow \neg Qx]$, $\forall x[Px \rightarrow Rx]$, and $\forall x[Px \rightarrow \neg Rx]$. Applying *Grice* to these alternatives means (among others) that the speaker knows none of them. In formulae, this means that $\neg\mathbf{K}\forall x[Px \rightarrow Qx]$, $\neg\mathbf{K}\forall x[Px \rightarrow \neg Qx]$, $\neg\mathbf{K}\forall x[Px \rightarrow Rx]$, and $\neg\mathbf{K}\forall x[Px \rightarrow \neg Rx]$. This is equivalent to saying that the following formulae are true: $\mathbf{P}\exists x[Px \wedge \neg Qx]$, $\mathbf{P}\exists x[Px \wedge Qx]$, $\mathbf{P}\exists x[Px \wedge \neg Rx]$, and $\mathbf{P}\exists x[Px \wedge Rx]$. To strengthen this inference, we apply competence again. The relevant notion of competence now, of course, is that the speaker knows which P -individuals have property Q and/or R . Making use of this competence assumption we can strengthen the possibility statements into knowledge attributions: $\mathbf{K}\exists x[Px \wedge \neg Qx]$, $\mathbf{K}\exists x[Px \wedge Qx]$, $\mathbf{K}\exists x[Px \wedge \neg Rx]$ and $\mathbf{K}\exists x[Px \wedge Rx]$. Because knowledge entails truth we infer (among others) to the conjunctive reading: $\exists x[Px \wedge Qx]$ and $\exists x[Px \wedge Rx]$.

is true and one where the speaker knows that only $\diamond_j P(b)$ is true. But – as we have seen in the beginning of this section – ‘only knowing’ doesn’t make sense in case we have more than one minimal state, so something has to be done. Perhaps what we do is to pragmatically reinterpret the sentence by first eliminating the above minimal states from the set of information states that we take as input (because these states can be expressed more economically by alternative expressions with a stronger meaning), and then apply the pragmatic interpretation function *Grice* to this more reduced set. In that case we end up with the desired result: we again have a unique minimal state, and in this state the speaker knows neither $\diamond_j P(c)$ nor $\diamond_j (P(a) \wedge P(b))$, but she does know both $\diamond_j P(a)$ and $\diamond_j P(b)$.

But, you will wonder, why does this procedure not also work for ‘ $p \vee q$ ’? Does this procedure not predict that from ‘ $p \vee q$ ’ we can conclude that the speaker knows both p and q ? Indeed, on assuming competence, also these examples give rise to two minimal states. Eliminating those states without giving up competence now results in ‘ $\mathbf{K}(p \wedge q)$ ’. For non-embedded disjunctions, however, we assume with others that their conjunctive counterparts are alternative expressions. But that means that in contrast to $\diamond_j (Pa \vee Pb)$, conveying such information *could* be done more transparently by semantically stronger alternative expressions, so in these cases, these interpretations are not allowed (notice that we use here some type of bidirectional interpretation procedure, taking also into account how the speaker would have expressed his information state).³⁰ Giving up competence is the only rescue now, which is possible for ‘ $p \vee q$ ’.

References

- [1] Burge, T. (1979), “Individualism and the mental, in P. French et al. (eds), *Midwest Studies in Philosophy, 4, Studies in Epistemology*, University of Minnesota Press, Minneapolis, pp. 73-122.
- [2] Chierchia, G. and S. McConnell-Ginet (1990), *Meaning and Grammar. An Introduction to Semantics*, MIT Press, Cambridge Massachusetts.
- [3] Donnellan, K. (1966), “Reference and definite descriptions”, *Philosophical Review*, 75, pp. 281-304.

³⁰For more explicit bidirectional analyses of free choice permissions, see Franke (2010) and van Rooij (2010).

- [4] Evans, G. (1973), “The causal theory of names”, *Proceedings of the Aristotelian Society*, Supplementary Volume 47, pp. 84-104.
- [5] Franke, M. (2010), ‘Free choice from iterated best response’, in M. Aloni & K. Schulz (eds.), *Proceedings of the Amsterdam Colloquium*.
- [6] Frege, G. (1892), “Über Sinn und Bedeutung”, *Zeitschrift für Philosophie und philosophische Kritik*, 50, pp. 25-50.
- [7] Gärdenfors, P. (2000), *Conceptual Spaces. The Geometry of Thought*, MIT Press, Cambridge, MA.
- [8] Gazdar, G. (1979), *Pragmatics*, Academic Press, London.
- [9] Grice, H.P. (1957), ‘Meaning’, *Philosophical Review*, 66: 377-88.
- [10] Grice, H. P. (1967), ‘Logic and Conversation’, typescript from the William James Lectures, Harvard University. Published in P. Grice (1989), *Studies in the Way of Worlds*, Harvard University Press, Cambridge Massachusetts, 22-40.
- [11] Groenendijk, J. and M. Stokhof (1984), *Studies in the Semantics of Questions and the Pragmatics of Answers*, Ph.D. thesis, University of Amsterdam.
- [12] Groenendijk, J. and M. Stokhof (1991), ‘Dynamic predicate logic’. *Linguistics and Philosophy* 14: 39-100.
- [13] Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. dissertation, University of Massachusetts, Amherst.
- [14] Heim, I. and A. Kratzer (1998), *Semantics in Generative Grammar*, Blackwell Publishers, Oxford.
- [15] Horn, L. (1972), *The semantics of logical operators in English*, Ph.D. thesis, Yale University.
- [16] Kamp, H. (1973), ‘Free choice permission’, *Proceedings of the Aristotelian Society*. N.S., 74: 57-74.
- [17] Kamp, H. (1979), ‘Semantics versus pragmatics’. In: F. Guenther and J. Schmidt (eds.), *Formal Semantics and Pragmatics for Natural Language*, Reidel, Dordrecht, 225-78.

- [18] Kaplan, D. (1989), “Demonstratives”, in I. Almog et al. (eds.) *Themes from Kaplan*, Oxford University Press, New York, pp. 481-563.
- [19] Katz, J.J. and P.M. Postal (1964), *An Integrated Theory of Linguistic Descriptions*, Cambridge, Mass.
- [20] Kripke, S. (1972/80), “Naming and necessity”, In: D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, Dordrecht, pp. 253-355, 763-769.
- [21] Lakoff, G. (1987), *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*, Chicago, University of Chicago Press.
- [22] Lewis, D. (1970/9), ‘A problem about permission’. In: E. Saarinen et al. (eds.), *Essays in Honour of Jaakko Hintikka*, Reidel, Dordrecht (ms. 1970).
- [23] Lewis, D. (1984) “Putnam’s paradox”, *The Australian Journal of Philosophy*, **62**: 221-236.
- [24] Millikan, R. (1984), *Language, Thought and Other Biological Categories*, MIT Press, Cambridge, MA.
- [25] Putnam, H. (1975), “The meaning of ‘meaning’”, In: K. Gunderson (ed.), *Language, Mind and Knowledge*, Minneapolis, MN, University of Minnesota Press.
- [26] Putnam, H. (1977), “Realism and reason”, *Proceedings of the American Philosophical Association*, **50**: 483-498.
- [27] Quine, W.V.O. (1960), *Word and Object*, Technology Press and John Wiley & Son, New York and London.
- [28] Rooij, R. van & K. Schulz (2004), ‘Exhaustive interpretation of complex sentences’, *Journal of Logic, Language and Information*.
- [29] Rooij, R. van (2010), ‘Conjunctive interpretation of disjunction’, *Semantics and Pragmatics*.
- [30] Rosch, E. (1978), ‘Principles of categorization’, In E. Rosch & B. Lloyd (eds.), *Cognition and categorization*, Hillsdale, NJ: Erlbaum.

- [31] Schulz, K. (2005), ‘A pragmatic solution to the paradox of free choice permission’. *Synthese*, **147**: 343-377.
- [32] Schulz, K. and R. van Rooij (to appear), ‘Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation’, *Linguistics and Philosophy*.
- [33] Searle, J.R. (1969), *Speech Acts*, Cambridge University Press, London and New York.
- [34] Spector, B. (2006), *Aspects de la pragmatique des opérateurs logiques*, PhD dissertation, University of Paris VII.
- [35] Stalnaker, R.C. (1974), “Pragmatic presupposition”, In: Munitz and Unger (eds.), *Semantics and Philosophy*, NYP.
- [36] Stalnaker, R.C. (1978), “Assertion”, In: P. Cole (ed.), *Syntax and Semantics, vol. 9: Pragmatics*, pp. 315-332.
- [37] Stalnaker, R. C. (1997), “Reference and necessity”, In: B. Hale and C. Wright (eds.), *A Companion to the Philosophy of Language*, Blackwell, Oxford.
- [38] Stalnaker, R.C. (2001), ‘On considering a possible world as actual’, *Proceedings of the Aristotelian Society*.
- [39] Strawson, P. F. (1950), “On referring”, *Mind*.
- [40] Veltman, Frank (1996), ‘Defaults in update semantics’. *Journal of Philosophical Logic*, **25**: 221-261.
- [41] Wittgenstein, L. (1953), *Philosophical Investigations*, Blackwell, Oxford.