5

Assertion
Robert C. Stalnaker

Let me begin with some truisms about assertions. First, assertions have content; an act of assertion is, among other things, the expression of a proposition – something that represents the world as being a certain way. Second, assertions are made in a context – a situation that includes a speaker with certain beliefs and intentions, and some people with their own beliefs and intentions to whom the assertion is addressed. Third, sometimes the content of the assertion is dependent on the context in which it is made, for example, on who is speaking or when the act of assertion takes place. Fourth, acts of assertion affect, and are intended to affect, the context, in particular the attitudes of the participants in the situation; how the assertion affects the context will depend on its content.

My aim in this paper is to sketch some theoretical concepts with which to develop these truisms, and to show how these concepts can be used to explain some linguistic phenomena. I want to suggest how content and context might be represented in a theory of speech, and how the interaction of content and context to which the above mentioned truisms point might be described. I will not propose an analysis of assertion, but I will make some modest claims about the way assertions act on the contexts in which they are made, and the way contexts constrain the interpretation of assertions. In conclusion, I will look briefly at an example of a phenomenon which I think these modest claims help to explain.

Three notions will play a central role in the theory I will sketch: the notion of a **proposition**, the notion of a **propositional concept**, and the notion of **speaker presupposition**. Each of these three notions will be defined or explained in terms of the notion of a **possible world**, or a possible state of the world, so one might think it important to begin with the question, what is a possible world? This is a good question, but I will not try to answer it here, and I am not sure that an abstract theory of speech should say very much in answer to it. In particular inquiries, deliberations, and conversations, alternative states of the subject matter in question are conceived in various different ways depending on the interests and attitudes of the participants in those activities. But one thing that is common to all such activities, and essential to them, is that the participants do seek to distinguish among alternative ways that things

might be, or might have been. It may be that the best way to bring out the formal structure of such activities is to focus on what is done with a given relevant set of alternative states of the world, setting aside questions about the nature of the alternatives themselves. The decision to treat possible worlds, or possible situations, as **primitive** elements in a theory of propositions and propositional attitudes does not require an ontological commitment to possible worlds as basic entities of the universe. Rather, it is a decision to theorize at a certain level of abstraction.[1]

The analysis of proposition in terms of possible worlds was first proposed in the context of intuitive semantics for modal logic.[2] The analysis is this: A proposition is a function from possible worlds into truth values (true or false). More roughly and intuitively, a proposition is a rule for determining a truth value as a function of the facts – of the way the world is. Or, a proposition is a way – any way – of picking out a set of possible states of affairs – all those for which the proposition takes the value true.

The intuitive motivation for this analysis is something like the following. A proposition – the content of an assertion or belief – is a representation of the world as being a certain way. But for any given representation of the world as being a certain way, there will be a set of all the possible states of the world which accord with the representation – which **are** that way. So any proposition determines a set of possible worlds. And, for any given set of possible worlds, to locate the actual world in that set is to represent the world as being a certain way. So every set of possible worlds determines a proposition. Furthermore, any two assertions or beliefs will represent the world as being the **same** way if and only if they are true in all the same possible worlds. If we assume, as seems reasonable, that representations which represent the world as being the same way have the same content (express the same proposition), then we can conclude that there is a one-one correspondence between sets of possible worlds and propositions. Given this correspondence, it seems reasonable to use sets of possible worlds or (equivalently) functions from possible worlds into truth values, to play the role of propositions in our theory. The analysis defines propositions in terms of their essential function – to represent the world.[3]

Supposing for convenience of exposition that there is just a small finite number of possible states of the world, we might represent a proposition by enumerating the truth values that it has in the different possible worlds, as in the following matrix:

$$
\begin{array}{c|ccc}
A & i & j & k \\
\hline
 & \text{T} & \text{F} & \text{T} \\
\end{array}
$$

i, j, and k are the possible worlds – the different possible sets of facts that determine the truth value of the proposition.

But there is also a second way that the facts enter into the determination of the truth value of what is expressed in an utterance: It is a matter of fact that an utterance has the content which it has. What one says – the proposition he expresses – is itself something that might have been different if the facts had been different; and if one is mistaken about the truth value of an utterance, this is sometimes to be explained as a misunderstanding of what was said rather than as a mistake about the truth value of what was actually said. The difference between the two ways that truth values depend on facts is exploited in the familiar riddle, *If you call a horse's tail a leg how many legs does a horse*

have? The answer, of course, is four, since calling a tail a leg does not make it one, but one can see a different way to take the question.

Let me give a simple example: I said *You are a fool* to O'Leary. O'Leary is a fool, so what I said was true, although O'Leary does not think so. Now Daniels, who is no fool and who knows it, was standing nearby, and he thought I was talking to him. So both O'Leary and Daniels thought I said something false: O'Leary understood what I said, but disagrees with me about the facts; Daniels, on the other hand, agrees with me about the fact (he knows that O'Leary is a fool), but misunderstood what I said. Just to fill out the example, let me add that O'Leary believes falsely that Daniels is a fool. Now compare the possible worlds i, j, and k. i is the world as it is, the world we are in; j is the world that O'Leary thinks we are in; and k is the world Daniels thinks we are in. If we ignore possible worlds other than i, j, and k, we can use matrix A to represent the proposition I actually expressed. But the following **two-dimensional** matrix also represents the second way that the truth value of my utterance is a function of the facts:

B	i	j	k
i	T	F	T
j	T	F	T
k	F	T	F

The vertical axis represents possible worlds in their role as context – as what determines what is said. The horizontal axis represents possible worlds in their role as the arguments of the functions which are the propositions expressed. Thus the different horizontal lines represent **what is said** in the utterance in various different possible contexts. Notice that the horizontal line following i is the same as the one following j. This represents the fact that O'Leary and I agree about what was said. Notice also that the vertical column under i is the same as the one under k. This represents the fact that Daniels and I agree about the truth values of both the proposition I in fact expressed and the one Daniels thought I expressed.

In a sense, I said something true at i and false at j and k, even though in none of these worlds did I express the proposition that is true in i and false in j and k. Although not expressed in any of the contexts, this proposition is represented in the matrix. I will call it the **diagonal proposition** since it is the function from possible worlds into truth values whose values are read along the diagonal of the matrix from upper left to lower right. In general, this is the proposition that is true at i for any i if and only if what is expressed in the utterance at i is true at i. I shall say more about diagonal propositions later.

I will call what a matrix like B represents a **propositional concept**. A propositional concept is a function from possible worlds into propositions, or, equivalently, a function from an ordered pair of possible worlds into a truth value. Each concrete utterance token can be associated with the propositional concept it determines, and, I will suggest below, some of the principles constraining the interpretation and evaluation of assertions are constraints on propositional concepts determined by assertive utterances rather than simply on the propositions expressed. This is my motivation for introducing propositional concepts, but one can study this kind of structure from an abstract point of view, independently of utterances or contexts of utterance. The

abstract theory of what I am calling propositional concepts has received some attention from logicians recently under the name **two-dimensional modal logic**.[4] The theory focusses on the notion of a two-dimensional modal operator.

A two-dimensional modal operator is an operator which takes a propositional concept into a propositional concept. If o is such an operator, then the meaning of o will be a rule that gives you the propositional concept expressed by oP in terms of the one expressed by P, for any P. I will describe one such operator, and contrast it with more traditional extensional and intensional sentence operators.[5]

The dagger is an operator which takes the diagonal proposition and projects it onto the horizontal. If ϕ is the diagonal proposition determined by P, then $\dagger P$ expresses ϕ relative to all contexts. So if B is the propositional concept determined by my statement to O'Leary in the example above, the following matrix gives the propositional concept, $\dagger B$:

$\dagger B$	i	j	k
i	T	F	F
j	T	F	F
k	T	F	F

What $\dagger B$ says is roughly this: *What is said in S's utterance of* **You are a fool** *is true*, where the definite description, *What is said in S's utterance of* **You are a fool** may be a nonrigid designator – a description that refers to different propositions in different worlds. Notice that the dagger always yields a constant propositional concept as its value. That is, whatever the case with P, $\dagger P$ will always express the same proposition relative to every context. If P itself is already a constant propositional concept in this sense, then $\dagger P$ will express the same propositional concept as P.[6]

Compare this operator with a more familiar modal operator, propositional necessity. $\Box P$ expresses in any world the proposition that is true at that world if and only if the proposition expressed by P at that world is the necessary proposition – the one that is true in all possible worlds. Propositional necessity is a one-dimensional operator in the following sense: The proposition expressed by $\Box P$ at any point depends only on the proposition expressed by P at that point. To evaluate $\Box P$ on any horizontal line, one need look only at the values of P on that line. This distinction between one-and two-dimensional operators parallels, on the next level up, the distinction between extensional and intensional operators. Compare the extensional negation operator: to evaluate $\sim P$ at any point, one need look only at the value of P at that point. Extensional operators take points (truth values) into points; one-dimensional operators take horizontal lines (propositions) into horizontal lines; two-dimensional operators take the whole matrix (the propositional concept) into another whole matrix. Each kind of operator is a generalization of the kind preceding it.[7]

Let me mention one complex operator, square-dagger, which says that the diagonal proposition is necessary. This can be understood as the **a priori truth** operator, observing the distinction emphasized in the work of Saul Kripke between a priori and necessary truth. An a priori truth is a statement that, while perhaps not expressing a necessary proposition, expresses a truth in every context. This will be the case if and only if the diagonal proposition is necessary, which is what the complex operator says.

I will illustrate this with a version of one of Kripke's own examples (1971: 273–5). Suppose that in worlds i, j, and k, a certain object, a metal bar, is one, two, and three meters long, respectively, at a certain time t. Now suppose an appropriate authority fixes the reference of the expression *one meter* by making the following statement in each of the worlds i, j, and k: *This bar is one meter long*. Matrix C below represents the propositional concept for this statement. Matrix $\square \dagger C$ represents the propositional concept for the claim that this statement is a priori true:

C	i	j	k
i	T	F	F
j	F	T	F
k	F	F	T

$\square \dagger C$	i	j	k
i	T	T	T
j	T	T	T
k	T	T	T

The proposition expressed by the authority is one that might have been false, although he couldn't have expressed a false proposition in that utterance.

I have said how propositions are to be understood, and what propositional concepts are. The third notion I need is the concept of speaker presupposition. This, I want to suggest, is the central concept needed to characterize speech contexts. Roughly speaking, the presuppositions of a speaker are the propositions whose truth he takes for granted as part of the background of the conversation. A proposition is presupposed if the speaker is disposed to act as if he assumes or believes that the proposition is true, and as if he assumes or believes that his audience assumes or believes that it is true as well. Presuppositions are what is taken by the speaker to be the **common ground** of the participants in the conversation, what is treated as their **common knowledge** or **mutual knowledge**.[8] The propositions presupposed in the intended sense need not really be common or mutual knowledge; the speaker need not even believe them. He may presuppose any proposition that he finds it convenient to assume for the purpose of the conversation, provided he is prepared to assume that his audience will assume it along with him.

It is **propositions** that are presupposed – functions from possible worlds into truth values. But the more fundamental way of representing the speaker's presuppositions is not as a set of propositions, but rather as a set of possible worlds, the possible worlds compatible with what is presupposed. This set, which I will call the **context set**, is the set of possible worlds recognized by the speaker to be the "live options" relevant to the conversation. A proposition is presupposed if and only if it is true in all of these possible worlds. The motivation for representing the speaker's presuppositions in terms of a set of possible worlds in this way is that this representation is appropriate to a description of the conversational process in terms of its essential purposes. To engage in conversation is, essentially, to distinguish among alternative possible ways that things may be. The purpose of expressing propositions is to make such distinctions. The presuppositions define the limits of the set of alternative possibilities among which speakers intend their expressions of propositions to distinguish.

Each participant in a conversation has his own context set, but it is part of the concept of presupposition that a speaker assumes that the members of his audience presuppose everything that he presupposes. We may define a **nondefective context** as one in which the presuppositions of the various participants in the conversation are all

the same. A **defective context** will have a kind of instability, and will tend to adjust to the equilibrium position of a nondefective context. Because hearers will interpret the purposes and content of what is said in terms of their own presuppositions, any unnoticed discrepancies between the presuppositions of speaker and addressees is likely to lead to a failure of communication. Since communication is the point of the enterprise, everyone will have a motive to try to keep the presuppositions the same. And because in the course of a conversation many clues are dropped about what is presupposed, participants will normally be able to tell that divergences exist if they do. So it is not unreasonable, I think, to assume that in the normal case contexts are nondefective, or at least close enough to being nondefective.

A context is **close enough** to being nondefective if the divergences do not affect the issues that actually arise in the course of the conversation. Suppose for example that you know that Jones won the election, believe mistakenly that I know it as well, and are prepared to take the truth of this proposition for granted if the occasion should arise, say by using it as a suppressed premise in an argument, or by using the description *the man who won the election* to refer to Jones. On my dispositional account of speaker presupposition, if you are prepared to use the proposition in this way, then you *do* presuppose that Jones won the election, even if you never have the opportunity to display this disposition because the subject does not come up. Since I do not know that Jones won the election, I do *not* presuppose it, and so the context is defective. But the defect may be harmless.

It will not necessarily be harmless: If the news is of sufficiently urgent interest, your failure to raise the subject may count as a display of your disposition to take its truth for granted. There will not be exactly a failure of communication, but there will be a misperception of the situation if I infer from the fact that you do not tell me who won that you do not know either.

A conversation is a process taking place in an ever-changing context. Think of a state of a context at any given moment as defined by the presuppositions of the participants as represented by their context sets. In the normal, nondefective case, the context sets will all be the same, so for this case we can talk of the context set of the conversation. Now how does an assertion change the context? There are two ways, the second of which, I will suggest, should be an essential component of the analysis of assertion. I will mention the first just to set it apart from the second: The fact that a speaker is speaking, saying the words he is saying in the way he is saying them, is a fact that is usually accessible to everyone present. Such observed facts can be expected to change the presumed common background knowledge of the speaker and his audience in the same way that any obviously observable change in the physical surroundings of the conversation will change the presumed common knowledge. If a goat walked into the room, it would normally be presupposed, from that point, that there was a goat in the room. And the fact that this was presupposed might be exploited in the conversation, as when someone asks, *How did that thing get in here?*, assuming that others will know what he is talking about. In the same way, when I speak, I presuppose that others know I am speaking, even if I do not assume that anyone knew I was going to speak before I did. This fact, too, can be exploited in the conversation, as when Daniels says *I am bald*, taking it for granted that his audience can figure out who is being said to be bald.

I mention this commonplace way that assertions change the context in order to make clear that the context on which an assertion has its **essential** effect is not defined by what is presupposed before the speaker begins to speak, but will include any information which the speaker assumes his audience can infer from the performance of the speech act.

Once the context is adjusted to accommodate the information that the particular utterance was produced, how does the **content** of an assertion alter the context? My suggestion is a very simple one: To make an assertion is to reduce the context set in a particular way, provided that there are no objections from the other participants in the conversation. The particular way in which the context set is reduced is that all of the possible situations incompatible with what is said are eliminated. To put it a slightly different way, the essential effect of an assertion is to change the presuppositions of the participants in the conversation by adding the content of what is asserted to what is presupposed. This effect is avoided only if the assertion is rejected.

I should emphasize that I do not propose this as a **definition** of assertion, but only as a claim about one effect which assertions have, and are intended to have – an effect that should be a component, or a consequence, of an adequate definition. There are several reasons why one cannot define assertion in terms of this effect alone. One reason is that other speech acts, like making suppositions, have and are intended to have the same effect. A second reason is that there may be various indirect, even nonlinguistic, means of accomplishing the same effect which I would not want to call assertions. A third reason is that the proposed essential effect makes reference to another speech act – the rejection of an assertion,[9] which presumably cannot be explained independently of assertion.

Our proposed effect is clearly not a sufficient condition for assertion. Is it even a necessary condition? It might be objected that a person who makes an assertion does not necessarily intend to get his audience to accept that what he asserts is true. The objector might argue as follows: Take one of your own examples, your statement to O'Leary that he is a fool. You knew in advance that O'Leary would not accept the assertion, so according to your account, you knew in advance that your assertion would fail to achieve its essential effect. That example should be anomalous if your account were correct, but it is not anomalous. Would it not be more plausible to characterize assertion as trying to get the audience to accept **that the speaker accepts** the content of the assertion?[10] But this Gricean twist is not required. My suggestion about the essential effect of assertion does not imply that speakers **intend** to succeed in getting the addressee to accept the content of the assertion, or that they believe they will, or even might succeed. A person may make an assertion knowing it will be rejected just as Congress may pass a law knowing it will be vetoed, a labor negotiator may make a proposal knowing it will be met by a counterproposal, or a poker player may place a bet knowing it will cause all the other players to fold. Such actions need not be pointless, since they all have secondary effects, and there is no reason why achieving the secondary effects cannot be the primary intention of the agent performing the action. The essential effects will still be relevant even when it is a foregone conclusion that the assertion, legislative act, proposal, or bet will be rejected, since one generally explains why the action has the secondary effects it has partly in terms of the fact that it would have had certain essential effects had it not been rejected.

One may think of a nondefective conversation as a game where the common context set is the playing field and the moves are either attempts to reduce the size of the set in certain ways or rejections of such moves by others. The participants have a common interest in reducing the size of the set, but their interests may diverge when it comes to the question of how it should be reduced. The overall point of the game will of course depend on what kind of conversation it is – for example, whether it is an exchange of information, an argument, or a briefing.

The game could be expanded by introducing other kinds of moves like making stipulations, temporary assumptions, or promises, asking questions, and giving commands and permissions.[11] Each of these kinds of linguistic action is presumably performed against a background of presuppositions, and can be understood partly in terms of the effect that it has, or is intended to have, on the presuppositions, and on the subsequent behavior, of the other participants in the conversation.

This is a very abstract, and a very simple, sketch of what goes on when someone says something to someone else. But there is enough in it to motivate some principles that are useful for explaining regularities of linguistic usage. I will mention three such rules which illustrate the interaction of context and content. Given the framework of propositions, presupposition, and assertion, the principles are all pretty obvious, which is as it should be. They are not intended as empirical generalizations about how particular languages or idiosyncratic social practices work. Rather, they are proposed as principles that can be defended as essential conditions of rational communication, as principles to which any rational agent would conform if he were engaged in a practice that fits the kind of very abstract and schematic sketch of communication that I have given.[12]

I will list the three principles and then discuss them in turn.

1   A proposition asserted is always true in some but not all of the possible worlds in the context set.
2   Any assertive utterance should express a proposition, relative to each possible world in the context set, and that proposition should have a truth value in each possible world in the context set.
3   The same proposition is expressed relative to each possible world in the context set.

The first principle says that a speaker should not assert what he presupposes to be true, or what he presupposes to be false. Given the meaning of presupposition and the essential effect ascribed to the act of assertion, this should be clear. To assert something incompatible with what is presupposed is self-defeating; one wants to reduce the context set, but not to eliminate it altogether. And to assert something which is already presupposed is to attempt to do something that is already done.

This rule, like the others, can be applied in several ways. If one could fix independently what was presupposed and what was said on a given occasion, then one could use the rule to evaluate the speaker's action. If he failed to conform to the rule, then he did something that, from the point of view of the conversation, was unreasonable, inefficient, disorderly, or uncooperative. But one can also use the rule, or the presumption that the speaker is conforming to the rule, as evidence of what was presupposed, or of what was said. Perhaps as more than just evidence. The rules may be taken to

define partially what is presupposed and what is said in a context by constraining the relation between them. So, if a speaker says something that admits of two interpretations, one compatible with the context set and one not, then the context, through the principle, disambiguates. If the speaker says something that seems prima facie to be trivial, one may take it as a clue that the speaker's context set is larger than was realized – that the context was defective – or one may look for another interpretation of what he said. There are thus three ways to react to an apparent violation of the rule: First, one may conclude that the context is not as it seems. Second, one may conclude that the speaker didn't say (or didn't mean) what he seemed to say (or to mean). Third, one may conclude that the rule was indeed violated. Since there is usually a lot of flexibility in both the context and the interpretation of what is said, the third reaction will be an unusual one, although it will not be unusual to use the rule to explain why some utterance would have been deviant if it had occurred in a given context.

The second principle concerns truth value gaps, and connects semantic presupposition with pragmatic speaker presupposition. The principle implies that if a sentence x semantically presupposes a proposition ϕ (in the sense that x expresses a truth or a falsehood only if ϕ is true), then ϕ is presupposed by the speaker in the sense of presupposition discussed above.

There are two different ways that a truth value gap may arise: a sentence may fail to express a proposition at all in some possible situation, or it may succeed in expressing a proposition, but express one that is a **partial** function – one that is undefined for certain possible worlds. Both kinds of truth value gap are excluded from the context set by this rule.

The rationale for this rule is as follows: The point of an assertion is to reduce the context set in a certain determinate way. But if the proposition is not true or false at some possible world, then it would be unclear whether that possible world is to be included in the reduced set or not. So the intentions of the speaker will be unclear.

Again this principle can be used in any of the three ways: to interpret what is said, as a clue to what is presupposed, or as a basis for evaluating the action of a speaker.

The third principle, which says that an utterance must express the **same** proposition relative to each possible world in the context set, is closely related in its motivation to a fundamental assumption of the logical atomists and the logical empiricist tradition. In Wittgenstein's terminology the assumption is this: Whether a proposition (read: sentence) has sense cannot depend on whether another proposition is true (cf. *Tractatus*; Proposition 2.0211). Meaning and truth must be sharply divided, according to this tradition, in order that one be able to use language to communicate in a determinate way. One must be able to tell what a statement says independently of any facts that might be relevant to determining its truth. Now it has always been clear that this kind of principle requires qualification, since it is a matter of fact that words mean what they mean. And the phenomena of context dependence are evidence of other ways in which what is said is a function of what is true. The framework of presupposition and assertion at once provides a natural way to qualify this traditional assumption so as to make it compatible with the phenomena, and a clear explanation of why it must hold in the qualified version. To see why the

principle must hold, look at the matrix for the propositional concept D. Suppose the context set consists of i, j, and k, and the speaker's utterance determines D. What would he be asking his audience to do? Something like this: If we are in world i, leave

D	i	j	k
i	T	T	T
j	F	F	T
k	F	T	T

the context set the same; if we are in world j, throw out worlds i and j, and if we are in world k, throw out just world i. But of course the audience does not know which of those worlds we are in, and if it did the assertion would be pointless. So the statement, made in that context, expresses an intention that is essentially ambiguous. Notice that the problem is not that the speaker's utterance has failed to determine a unique proposition. Assuming that one of the worlds i, j, or k is in fact the actual world, then that world will fix the proposition unambiguously. The problem is that since it is unknown which proposition it is that is expressed, the expression of it cannot do the job that it is supposed to do.[13]

 As with the other principles, one may respond to apparent violations in different ways. One could take an apparent violation as evidence that the speaker's context set was smaller than it was thought to be, and eliminate possible worlds relative to which the utterance receives a divergent interpretation. Or, one could reinterpret the utterance so that it expresses the same proposition in each possible world. Consider an example: hearing a woman talking in the next room, I tell you, *That is either Zsa Zsa Gabor or Elizabeth Anscombe*. Assuming that both demonstrative pronouns and proper names are rigid designators – terms that refer to the same individual in all possible worlds – this sentence comes out expressing either a necessary truth or a necessary falsehood, depending on whether it is one of the two mentioned women or someone else who is in the next room. Let i be the world in which it is Miss Gabor, j the world in which it is Professor Anscombe, and k a world in which it is someone else, say Tricia Nixon Cox. Now if we try to bring the initial context set into conformity with the third principle by shrinking it, say by throwing out world k, we will bring it into conflict with the first principle by making the assertion trivial. But if we look at what is actually going on in the example, if we ask what possible states of affairs the speaker would be trying to exclude from the context set if he made that statement, we can work backward to the proposition expressed. A moment's reflection shows that what the speaker is saying is that the actual world is either i or j, and not k. What he means to communicate is that the diagonal proposition of the matrix E exhibited below, the proposition expressed by $\dagger E$, is true.

E	i	j	k		$\dagger E$	i	j	k
i	T	T	T		i	T	T	F
j	T	T	T		j	T	T	F
k	F	F	F		k	T	T	F

I suggest that a common way of bringing utterances into conformity with the third principle is to interpret them to express the diagonal proposition, or to perform on them the operation represented by the two-dimensional operator **dagger**. There are lots of examples. Consider: *Hesperus is identical with Phosphorus*, *it is now three o'clock*, *an ophthalmologist is an eye doctor*. In each case, to construct a context which conforms to the first principle, a context in which the proposition expressed is neither trivial nor assumed false, one must include possible worlds in which the sentence, interpreted in the standard way, expresses different propositions. But in any plausible context in which one of these sentences might reasonably be used, it is clear that the diagonal proposition is the one that the speaker means to communicate. The two-dimensional operator **dagger** may represent a common operation used to interpret, or reinterpret, assertions and other speech acts so as to bring them into conformity with the third principle constraining acts of assertion.

To conclude, let me show how this last suggestion can help to explain a puzzle concerning singular negative existential statements. The puzzle arises in the context of a causal or historical explanation theory of reference according to which proper names refer to their bearers, not in virtue of the fact that the bearer has certain properties expressed in the sense of the name, but rather in virtue of certain causal or historical connections between the referent and the speaker's use of the name.[14] According to this theory, the **proposition** expressed by a simple singular statement containing a proper name, like *O'Leary is a fool*, is the one that is true if and only if the individual who is in fact causally connected in the right way with the speaker's use of the name has the property expressed in the predicate. So the proposition is determined as a function of the individual named rather than as a function of the name, or the sense of the name.

What does this theory say about statements like *O'Leary does not exist?* If the statement is true (which this one happens to be), then there is no individual appropriately related to the speaker's use of the name, and thus no proposition determined as a function of such an individual. So at least for **true** negative existential statements, it seems that proper names must play a different role in the determination of the proposition expressed from the role they play in ordinary predicative statements.

Perhaps a negative existential statement says, simply, that there is no individual standing in the right causal relation to the speaker's use of the name.[15] This does seem to get the truth conditions right for negative existential **assertions**, but it clearly gets them wrong for some other kinds of singular negative existential constructions. Consider, for example, counterfactual suppositions, as in the antecedent of the conditional *If Aristotle hadn't existed, the history of philosophy would have been very different from the way it was*.[16] Clearly the proposition expressed in the antecedent of this conditional is not the proposition that our use of the name *Aristotle* is not appropriately connected with any individual. THAT proposition is compatible with Aristotle's existence. Furthermore, if Aristotle hadn't existed, then our uses of his name probably would not have existed either. The proper name seems to function in the antecedent of the counterfactual more like the way it functions in ordinary predicative statements: The proposition is determined as a function of the **person** Aristotle; it is true in possible worlds where **he** does not exist, and false in possible worlds where **he** does exist.

So it seems that not only do proper names act differently in negative existential assertions than they do in singular predicative assertions, they also act differently in

negative existential **assertions** than they do in negative existential **suppositions**. What one asserts when he says *Aristotle does not exist* seems to be different from what one supposes when he says *Suppose Aristotle hadn't existed*.

Let us see how the pragmatic principle can account for these facts. Begin with the most straightforward semantic account of negative existential constructions: *Aristotle does not exist*, like *Aristotle was wise*, is a proposition about Aristotle. It is true in possible worlds whose domains contain the person **we** call Aristotle and false in possible worlds whose domains do not contain that person. What if the name does not, in fact, refer? Suppose for example the statement is *Sherlock Holmes does not exist*. Then the proposition will be necessarily false, by the same rule, since the domain of no possible world contains the actual person **we** call Sherlock Holmes.[17] Now let us use this straightforward semantic account to construct a propositional concept for an utterance of *Sherlock Holmes does not exist*. Let the world i be the actual world. Let j be a world in which a famous detective named *Sherlock Holmes* lived in nineteenth century London, and Sir Arthur Conan Doyle wrote a series of historical accounts of his cases. Let world k be a possible world in which Sir Arthur Conan Doyle was a famous detective named *Sherlock Holmes* who wrote a series of autobiographical accounts of his own cases under the the pseudonym *Sir Arthur Conan Doyle*. These stipulations determine the following two-dimensional matrix for the utterance:

G	i	j	k
i	T	T	T
j	T	F	T
k	F	F	F

Now suppose i, j, and k are a context set (say a person has heard these three rumors about the origin of the Sherlock Holmes stories and does not know which is true). As the matrix shows, the utterance violates the third principle, and so a reinterpretation is forced on it. Diagonalization, or the dagger operation, brings the utterance into line with the principle, and yields the intuitively right result:

	i	j	k
i	T	F	F
j	T	F	F
k	T	F	F

But now contrast the case of the counterfactual. To interpret the statement *If Aristotle hadn't existed, the history of philosophy would have been very different from the way it was*, we do not need to diagonalize, since in any possible context appropriate to **that** statement, it will be presupposed that Aristotle does exist. So the proposition supposed is the one obtained by the straightforward rule.[18] Again, this is intuitively the right result.

We have not escaped the conclusion that the content of the assertion *Aristotle did not exist* is different from the content of the supposition *suppose Aristotle hadn't existed*. But we have explained that consequence using a single **semantic** account of singular

negative existential constructions – the account which is most natural, given the causal theory of names – together with independently motivated pragmatic principles.

The general strategy which this explanation illustrates is to use pragmatic theory – theory of conversational contexts – to take some of the weight off semantic and syntactic theory. Some other problems where I think this strategy and this theory will prove useful are the explanation of presupposition phenomena,[19] the explanation of the differences between subjunctive and indicative conditionals,[20] the analysis of definite descriptions, and the behavior of deictic and anaphoric pronouns. My hope is that by recognizing the interaction of some relatively simple contextual factors with the rules for interpreting and evaluating utterances, one can defend simpler semantic and grammatical analyses and give more natural explanations of many linguistic phenomena.

Notes

1   I argued in Stalnaker (1976a) that one can take possible worlds seriously without accepting an implausible metaphysics.
2   The possible worlds analysis of propositions was suggested originally by Saul Kripke in the early 1960s.
3   I recognize that I am skating quickly over large problems here. In particular, the identity conditions which the analysis assigns to propositions have some extremely paradoxical conse-quences (such as that there is only one necessary proposition) which seem to make the analysis particularly unsuited for an account of the objects of propositional attitudes. I discuss some of these problems, inconclusively, in Stalnaker (1976b).
4   The most general discussion of two-dimensional modal logic I know of is in Segerberg (1973). See also Aqvist (1973) and Kamp (1971). The earliest investigations of two-dimensional operators were, I believe, carried out in the context of tense logic by Frank Vlach and Hans Kamp at UCLA.
5   The tense logic analogue of the dagger operator was, according to David Lewis, invented by Frank Vlach and is discussed in his UCLA PhD dissertation (1974). The notation is Lewis's. See Lewis (1973a: 63–64*n*).
6   Another operator which has intuitive application is represented by Lewis as an upside-down dagger. What it does is to project the diagonal proposition onto the **vertical**, which, in effect, turns contingent truths into necessary truths and contingent falsehoods into necessary falsehoods. Hans Kamp (1971) proposed the temporal analogue of this operator as a representative of the sentence adverb *now*. *It is now true that A* said at time *t* expresses a proposition that is true at all times just in case *A* is true at *t*. The operator makes a difference when *now* is embedded in the context of other temporal modifiers. Using it, one can represent sentences like *Once, everyone now alive hadn't yet been born* without object language quantifiers over times. David Lewis and David Kaplan have suggested that this operator shows the semantic function of expressions like *actually* and *in fact*, as in *If I had more money than I in fact have, I would be happier*.
7   Although the dagger and the upside-down dagger are defined as operators on propositional concepts, they can be generalized to any kind of two-dimensional intension. For example, they

may be interpreted as operators on two-dimensional individual concepts, or on property concepts. Let *a* represent a definite description, say *the President of the United States*, and let *i*, *j*, and *k* be three times, say 1967, 1971, and 1975. Matrix (i) below represents the two-dimensional intension of this definite description relative to these times. Matrix (ii) represents the rigid description, *the person who is **in fact**, or **now**, the President of the United States*. This is the two-dimensional intension of $\downarrow a$. David Kaplan, in "DTHAT" (1978), discusses this operator on singular terms and compares it with Keith Donnellan's account of the referential use of definite descriptions.

(i)

	i	*j*	*k*
i	LJ	RN	GF
j	LJ	RN	GF
k	LJ	RN	GF

(ii)

	i	*j*	*k*
i	LJ	LJ	LJ
j	RN	RN	RN
k	GF	GF	GF

8   I have discussed this concept of presupposition in two earlier papers, Stalnaker (1973) and Stalnaker (1974). Stephen Schiffer (1972: 30–42) and David Lewis (1969: 52–60) have discussed concepts of mutual knowledge and common knowledge which resemble the notion of presupposition I have in mind. Paul Grice spoke, in the William James Lectures, of propositions having **common ground status** in a conversation (published in part in Grice (1975)).

9   It should be made clear that to reject an assertion is not to assert or assent to the contradictory of the assertion, but only to refuse to accept the assertion. If an assertion is rejected, the context remains as it was. (More exactly, rejection of an assertion blocks the **second** kind of effect that assertions have on the context. The first kind of effect cannot be blocked or withdrawn.)

10 David Kaplan, in discussion, raised this objection.

11 David Lewis (1973b) outlined a language game of commanding and permitting which would fit into this framework.

12 The influence of Grice's theory of conversation should be clear from my discussion of the application of these principles.

13 A clarification is needed to resolve an ambiguity. The third principle says that the proposition expressed in any possible world in the context set must coincide **within the context set** with the proposition expressed in any other possible world in the context set. So, for example, if the context set is $\{i, j\}$, then an utterance determining the propositional concept represented below will not violate the principle. Even though the proposition expressed in *i* diverges from the proposition expressed in *j*, the divergence is outside the context set. David Lewis pointed out the need for this clarification.

	i	*j*	*k*
i	T	F	T
j	T	F	F
k	F	T	T

14 The causal account of reference is defended, in general, in Kripke (1971) and Donnellan (1971) Donnellan (1974) discusses the problem of singular negative existential statements in the context of this account of reference.

15   Donnellan's explanation of the **truth conditions** for singular negative existential statements is roughly in accord with this suggestion, but he cautions that the rule he proposes "does not provide an **analysis** of such statements; it does not tell us what such statements mean, or what propositions they express. This means that in this case we are divorcing truth conditions from meaning" (1974: 25). According to Donnellan, "no obvious way of representing propositions expressed by existential statements suggests itself" (1974: 30).

16   Kripke, in talks on this subject, has made this point about counterfactuals with negative existential antecedents.

17   I believe this straightforward semantic account is the one that Kripke has defended in the talks mentioned in note 16.

18   It is interesting to note that if the conditional were in the indicative mood, the result would have been different. This is because an indicative conditional is appropriate only in a context where it is an open question whether the antecedent is true. So to say *If Aristotle didn't exist* is to suppose just what is asserted when one asserts *Aristotle didn't exist*.

19   This is discussed in Stalnaker (1973).

20   This is discussed in Stalnaker (1976c).

References

Aqvist, L. 1973. Modal logic with subjunctive conditionals and dispositional predicates. *Journal of Philosophical Logic* 2: 1–76.

Donnellan, K. 1971. Proper names and identifying descriptions. In D. Davidson and G. Harman (eds), *Semantics of Natural Language*, Dordrecht: D. Reidel.

Donnellan, K. 1974. Speaking of nothing. *Philosophical Review* 83: 3–31.

Grice, H. P. 1975. Logic and conversation. In P. Cole and J. H. Morgan (eds), *Speech Acts* (*Syntax and Semantics* vol. 3), New York: Academic Press.

Kamp, H. 1971. Formal properties of "Now". *Theoria* 37: 227–73.

Kaplan, David. 1978. Dthat. In Peter Cole (ed.), *Pragmatics* (*Syntax and Semantics*, vol. 9), New York: Academic Press.

Kripke, S. 1971. Naming and necessity. In D. Davidson and G. Harman (eds), *Semantics of Natural Language*, Dordrecht: D. Reidel.

Lewis, D. 1969. *Convention*. Cambridge, Mass.: Harvard University Press.

Lewis, D. 1973a. *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. 1973b. A Problem about Permission, unpublished paper, Princeton University, Princeton, New Jersey.

Schiffer, S. 1972. *Meaning*, Oxford: Clarendon Press.

Segerberg, K. 1973. Two-dimensional modal logic. *Journal of Philosophical Logic* 2: 77–96.

Stalnaker, R. C. 1973. Presuppositions. *Journal of Philosophical Logic* 2: 447–57.

Stalnaker, R. C. 1974. Pragmatic presuppositions. In M. K. Munitz and P. K. Unger (eds), *Semantics and Philosophy*, New York: New York University Press.

Stalnaker, R. C. 1976a. Possible worlds. *Nous* 10: 65–75.

Stalnaker, R. C. 1976b. Propositions. In A. F. Mackay and D. D. Merrill (eds), *Issues in the Philosophy of Language*, New Haven, Conn: Yale University Press.

Stalnaker, R. C. 1976c. Indicative conditionals. In A. Kasher (ed.), *Language in Focus: Foundations, Methods, and Systems*, essays in honor of Yehoshua Bar-Hillel, Dordrecht: D. Reidel.

Vlach, F. 1974. Unpublished Doctoral dissertation, University of California, Los Angeles, California.

Wittgenstein, L. 1961. *Tractatus Logico-Philosophicus* (Translation). New York: Routledge and Kegan Paul (Originally published in 1921).