POSSIBLE AND MUST

LAURI KARTTUNEN
University of Texas at Austin

This chapter is a limited account of the semantics of possibility and necessity in ordinary English. I will be concerned with words like possible, may, can, perhaps, and conceivable only to the extent they express epistemic or logical possibility. Whatever difference there may be between the examples in (1) remains outside the scope of the present study.

(1)  
a. It is possible that it is raining in Chicago.  
b. It may be raining in Chicago.  
c. Perhaps it is raining in Chicago.  
d. It is conceivable that it is raining in Chicago.  

I will discuss expressions like must, necessarily, have to, and is bound to only in their epistemic sense, that is, the sense in which they appear in the following set of examples:

(2)  
a. It must be raining in Chicago.  
b. It has to be raining in Chicago.  
c. Necessarily, it is raining in Chicago.  
d. It is bound to be raining in Chicago.

1 This work was supported in part by grant GS 2468 from the National Science Foundation. An earlier version of this paper was presented at the Seventh Regional Meeting of the Chicago Linguistic Society, April 1971. I am indebted to George Lakoff for suggesting a number of improvements in the manuscript.
I will not discuss the so-called deontic sense of *must* or the permissive use of *may* illustrated in (3):

(3)  
   a. *I must finish this talk in 20 minutes.*  
   b. *May the audience ask questions?*

In discussing the notions of possibility and necessity in ordinary language, I will employ Jaakko Hintikka's notions model set and model system (Hintikka, 1962, 1969). Since these concepts may not be familiar to all linguists, I will start with a brief introduction to these concepts.

**LOGICAL MODALITIES**

A model set, as defined by Hintikka, is a set of formulas that can be thought of as a partial description of a possible state of affairs ("possible world"). A model set has to meet certain basic conditions, such as those in (4):

(4)  
   a. Any model set $W$ that contains some atomic formula $p$ does not contain its negation; i.e., if $p \in W$, then $\neg p \notin W$.  
     \[ (C.\neg) \]  
   b. if $(p \& q) \in W$, then $p \in W$ and $q \in W$.  
     \[ (C.\&) \]  
   c. If $(p \lor q) \in W$, then $p \in W$ or $q \in W$.  
     \[ (C.\lor) \]

In Hintikka's system, the conditions in (4) obviously play the same role as truth tables play in standard treatments of sentential logic. In a complete system, one would also need corresponding conditions for negated formulas, identities, and quantifiers, but they will not be used in this paper.

Hintikka defines a model system $Q$ as a set of model sets related to each other by a relation of alternativeness. Intuitively, an alternative to a given model set $W$ is a partial description of a possible state of affairs that could have been realized instead of the one described by $W$. The basic modal concepts of necessity $L$ and possibility $M$ are captured in this framework by requiring that the following three conditions be satisfied by each model system $Q$ and its alternativeness relation:

(5)  
   a. If $Lp$ ("necessarily $p$") $\in W \in Q$, then $p \in W$.  
   b. If $Mp$ ("possibly $p$") $\in W \in Q$, then there is in $Q$ at least one alternative to $W$ that contains $p$.  
   c. If $Lp \in W \in Q$ and if $W^*$ is an alternative to $W$ in $Q$, then $p \in W^*$. 
In other words, what is possibly true in some state of affairs W is true in at least one of its alternatives. Whatever is necessarily true in W is true, not only in W itself, but in all of its alternatives as well.

There are a number of additional features that may be imposed on a model system by requiring that the alternativeness relation must also meet other constraints, such as transitivity, or symmetry. Each additional condition gives rise to a new system of modal logic. What is important to realize is that, in a sense, there is no “correct” way to construct a modal logic, no unique set of constraints on the alternativeness relation in a model set. Instead, there are a number of different systems, each of which is a legitimate object for study in its own right. Assuming that all of these systems share the same basic conditions [such as those in (4)], which define the notion of model set, they will not differ with respect to the interpretation of simple logical connectives. They are different only with respect to their modal operators, since the semantics of necessarily and possibly crucially depend on how the alternativeness relation is defined.

Assuming this framework, I shall now examine the way in which the words possibly, may, must, etc., are used in ordinary English. In particular, I shall investigate the sort of constraints that would have to be placed on the alternativeness relation in a model system that reflects the natural or naive modal concepts of ordinary language discourse. Not surprisingly, I will argue that the principles on which our use of possible and must in ordinary language appears to be based are different from any of the well known standard systems that have been studied by logicians. Of course, if I am right, this is an interesting fact about ordinary language, but it should not be misconstrued as evidence against other modal systems.

THE WORD POSSIBLE

There is an interesting constraint on the use of possible that we all seem to follow in our ordinary language conversation. This rule can be phrased simply as in (6).

(6) WHATEVER IS CANNOT POSSIBLY BE OTHERWISE.

To see this principle at work, consider examples like those in (7), which are deviant utterances in ordinary English:

(7) a. *It isn’t raining in Chicago, but it may be raining there.
    b. *Mary is pretty, but perhaps she is not pretty.
    c. *John is mistaken, but it is possible that he is right.
    d. *It is possible that God is alive, but he is dead.
One must be careful to note that the ill formedness of the above examples could be remedied very easily by a couple of small changes. For example, (7a) becomes perfectly acceptable if we prefix the first conjunct with *I think* or *I believe*.

\[(7a') \quad I \text{ think it isn't raining in Chicago, but it may be raining there.}\]

On the other hand, *I know* as a prefix preserves the anomaly, as shown in (7a’’):

\[(7a'') \quad I \text{ know it isn't raining in Chicago, but it may be raining there.}\]

I will return shortly to the difference between the *I think* and *I know* versions of (7a).

Another important feature in (7) is the mood of the verb. The ill formedness of the examples can be remedied by introducing a subjunctive conditional. Obviously, there is nothing wrong with the examples in (8):

\[(8) \quad \begin{array}{l}
\text{a. It isn't raining in Chicago, but it could be raining there if the clouds hadn't dispersed.} \\
\text{b. Mary is pretty, but it is possible that she wouldn't be pretty if she didn't wear any makeup.} 
\end{array}\]

Therefore, the principle in (6) must be carefully distinguished from a very similar sounding but incorrect principle (6’):

\[(6') \quad \text{Whatever is could not possibly be otherwise.}\]

The explicit presence of a subjunctive if clause in (8a) is not essential as long as we have *could* instead of *can*. However, it seems to me that examples such as (8a’) are to be interpreted as elliptic:

\[(8a') \quad I \text{t isn't raining in Chicago, but it could be.}\]

Some subjunctive conditional must be implicitly understood here. Example (8a’) says that it could be raining in Chicago, if something weren’t the way it is. The difference between *can* and *could, is possible* and *would be possible* will be discussed again below. For the time being, let us concentrate on the first type.

The anomaly of the examples in (7) depends on the fact that the two clauses do not differ with respect to temporal and locative references. This is shown in (9):
(9)  a. *God is alive, but tomorrow he may be dead.*
   b. *John is mistaken, but, under other circumstances, he might be right.*

What distinguishes the well formed examples in (8) and (9) from those in (7) is that the possibility clause makes reference to some other context than that in which the nonmodal sentence is supposed to be true.

Thus far, we have established that it is not acceptable in natural language to state in one sentence that something both is the case and possibly is not the case. The anomaly has nothing to do with syntactic well formedness in the usual sense. It is a matter of the meaning of *possible.* Consider a speaker who realizes that the rain is pouring down from the sky. He cannot honestly proclaim (10), unless he doubts his own senses:

(10)    *It is possible that it isn’t raining.*

At first it may seem that, under those circumstances, (10) is not any more deviant than (11):

(11)    *It is possible that it is raining.*

However, there is a difference. One can explain the unappropriateness of (11) on the basis of a general conversational postulate formulated by Grice (1968) that requires that in general one should make one’s statements as informative as one honestly can, provided that the information is relevant to the listener. In most cases, a man who is in the position to state simply *It is raining* would be extremely uncooperative if he chose to inform his listeners with (11) instead. He could certainly be accused of misleading, but not of lying. The difference shows up especially clearly in answers to the corresponding questions (12a) and (12b):

(12)    a. *Is it possible that it isn’t raining?*
   b. *Is it possible that it is raining?*

A man who knows that it is raining will have to say *No* to (12a) and *Yes* (or perhaps *Well, yes*) to (12b). [Since it is likely that whoever asks (12b) is interested in whether it in fact is raining, it is somewhat uncooperative to simply say *Yes* in case one actually knows that it is raining. By answering *Well, yes,* the speaker indicates explicitly that he is not giving directly the information that the questioner is seeking. On the use of *well,* see Robin Lakoff (1971).] That is, if it really is raining, then (10) is false and (11) is true. Note also that, although
the sentences in (13) could be called pointless or redundant, they are not incoherent in the same way as the examples in (7):²

(13)  ?It isn’t raining in Chicago; therefore, it is possible that it isn’t raining there.
     ?Mary is pretty, and it is possible that she is pretty.

POSSIBLE AND THE M-OPERATOR

Although the principle in (6) thus seems to hold for conversational English, it is not a valid principle in any of the standard modal logics.³ Corresponding to the deviant sentences in (7), we have the formula in (14), which is not inconsistent in any standard modal system.

(14)  \[ p \land M \sim p \]  \( p \) and it is possible that \( \neg p \)

There is no reason why (14) could not be a member of some model set. For example, assume that there is some model set \( W \) that contains (14). Condition (4b) requires that \( W \) must also contain the for-

² Aristotle distinguished between one-sided possibility and two-sided possibility (see Bochenski, 1961, pp. 81–83). In case of two-sided possibility, if \( p \) is possible, then \( \neg p \) is also possible. On the other hand, one-sided possibility obtains whenever something in fact is known to be the case. If something is necessary, then it is also (one-sidedly) possible. One could explain the strangeness of the examples in (13) by saying that they require the one-sided sense of possible, which is less natural in ordinary language than the two-sided reading. However, I doubt that there is any need to postulate these two distinct senses. As far as I understand it, Aristotle’s distinction is designed to account for the same facts that are also covered by Grice’s conversational postulates. Assuming that the speaker is following the cooperative principle by saying that something is possible, he indicates that he is not in the position to make a stronger statement; i.e., he does not know what the facts are. Therefore, for all he knows, the contrary is also possible. The two-sided interpretation of possible arises from these considerations; it is not part of the meaning of possible.

³ To my knowledge, the only modern logician who has proposed a system in which \( p \) and \( M \sim p \) are incompatible is J. Lukasiewicz (1920). His definition of possibility makes \( M \sim p \) true whenever \( p \) is either false or indeterminate. On the other hand, whenever \( p \) is true, \( M \sim p \) is false. It has been argued (see Kneale, 1962, p. 573) that Lukasiewicz’s three truth values (true, indeterminate, false) are not really truth values at all but certainty values. According to this interpretation, Lukasiewicz’s system can be translated into ordinary two-valued logic by reading “true” as “it is certain that,” “indeterminate” as “it is not certain that and it is not certain that not,” and “false” as “it is certain that not.” Furthermore, the Kneales suggest that “certain” can be taken as an abbreviation for “necessary in relation to what we know.” This interpretation of Lukasiewicz results in a system that is very similar if not identical to the one constructed by Jaakko Hintikka in his Knowledge and Belief. The modalities in this system are epistemic modalities, defined in terms of what we know.
formulas $p$ and $M \sim p$, as illustrated in (15). Furthermore, condition (5b) requires that there be some other model set, call it $W^\circ$, that is an alternative to $W$ and contains the formula $\neg p$:

\begin{align}
W &= \{ p & \land M \sim p, p, M \sim p \} \\
W^\circ &= \{ \neg p \}
\end{align}

The two model sets, $W$ and $W^\circ$, together form a model system that satisfies all the constraints given in (4) and (5). By constructing it, we have shown that (14) is not inconsistent.

Let us now look into the question of whether it would be feasible to add the principle.

(6) **Whatever is cannot be otherwise.**

to the set of conditions in (5), which define the notion of a model system. In this way we could perhaps obtain a new modal logic which more closely resembles natural languages than any of the standard logics. Formally, (6) would be expressed by a condition such as (16):

\begin{align}
\text{If } p &\in W \in Q, \text{ then } M \sim p \notin W \text{ [where } p \text{ is any nonmodal formula].}
\end{align}

In other words, whatever is true in some possible world is not false in any of its alternatives. An equivalent way to express the same principle is to state it as a condition on the alternativeness relation. This is given in (17):

(17) In each model set, the subset of nonmodal formulas (= whatever is) is included in all of the alternatives of that model set as well (= cannot be otherwise).

It is obvious that if the alternativeness relation between possible worlds is constrained by a condition like (17), the formula $p \land M \sim p$ cannot be a member of any model set, since there would have to be some alternative model set containing both $p$ and $\neg p$, which violates condition (4a). Thus, we would have a system that reflects the fact that corresponding English sentences are perceived as deviant. However, it is just as obvious that the resulting system could hardly be considered a modal logic at all. In the standard systems, the notions of necessity and possibility are interdefinable, as shown in (18):

(18) It is possible that not-$p \equiv$ it is not necessary that $p$. \\
$(M \sim p \equiv \neg Lp)$
Given the standard relation between logical necessity and possibility (16) is equivalent to (19):

(19) If $p \in W \in Q$, then $Lp \in W$ [where $p$ is any nonmodal formula].

In other words, whatever is true is necessarily true, if it cannot possibly be otherwise. Conditions (5a) and (19) together make $p$ logically equivalent with $Lp$; in effect, the notion of logical necessity simply collapses. In any standard axiomatic treatment of modal logic, the notion of possibility would consequently collapse as well. Here it does not happen except for the technical reason that model sets are defined as incomplete descriptions of possible worlds. That is, it is not required that for any formula $p$, a model set must contain either $p$ or its negation. From the fact that $Mp$ is contained in some model set, one cannot conclude from (16) that $p$ belongs to it as well, only that its negation does not. This line of investigation leads nowhere, except that it shows explicitly that there must be a basic mistake in thinking that the words may and possible in examples like (7a) represent the same notion of possibility that is studied in modal logic. What is involved here can be seen more clearly from the distinction between (7a') and (7a''). In saying something like It is possible that it is raining, the speaker states that the truth of it is raining is compatible with everything he knows about the world. If he knows that it is not raining, then the truth of it is possible that it is raining is not compatible with his knowledge. What we have here is an epistemic sense of possible.

EPISTEMIC LOGIC

This observation is not a new one. For example, Gottlob Frege [in his Begriffschrift in 1879] argued that modal distinctions always involve a covert reference to human knowledge; therefore, he thought that they had no place in pure logic. According to Frege, to say that something is possible is to say that the speaker knows nothing from which the negation of the proposition would follow.

More recently, Jaakko Hintikka (1962), explicitly defined epistemic possibility in terms of knowledge. In Hintikka's system, the two expressions in (20) are interchangeable:

(20) $a$ doesn't know that $p \equiv$ for all $a$ knows, it is possible that $\sim p$

($\sim K_a p \equiv P_a \sim p$)
In order to understand what Hintikka means by the expression \( \neg K_a p \), one has to keep in mind that, although Hintikka reads it as "a doesn’t know that \( p \),” he is using the verb know in a technical sense without the usual factive presupposition. If one does not want to do that much violence to natural language, \( \neg K_a p \) has to be read as suggested in (21):

\begin{equation}
\sim K_a p = \text{what } a \text{ knows is not that } p
\end{equation}

Hintikka shows how the semantics of epistemic statements can be stated in terms of conditions on model sets and alternativeness relations. The most important conditions are given in (22):

\begin{enumerate}
\item If \( K_a p \in W \), then \( p \in W \). \quad (C.K)
\item If \( \neg K_a p \in W \), then \( P_a \sim p \in W \). \quad (C, \neg K)
\item If \( P_a p \in W \) and if \( W \) belongs to a model system \( Q \), then there is in \( Q \) at least one alternative \( W^o \) to \( W \) (with respect to \( a \)) such that \( p \in W^o \). \quad (C,P^o)
\item If \( K_a q \in W \) and if \( W^o \) is an alternative to \( W \) (with respect to \( a \)) in some model system \( Q \), then \( K_a q \in W^o \). \quad (C, KK^o)
\end{enumerate}

Given the interdefinability of possibility and knowledge in Hintikka’s system, it follows directly that any expression of the type (7a’’) is inconsistent:

\begin{equation}
(7a’’) \quad \text{I know it isn’t raining in Chicago, but it may be raining there.}
\end{equation}

Example (7a’’) is represented by the formula in (23a), which immediately yields the contradiction in (23b) [see (20) above]:

\begin{enumerate}
\item \( K_a p \land P_a \sim p \)
\quad a knows that \( p \) and, for all that \( a \) knows, it is possible that \( \neg p \)
\item \( K_a p \land \neg K_a p \)
\quad a knows that \( p \) and \( a \) doesn’t know that \( p \)
\end{enumerate}

However, sentences of the type \( p \land P_a \sim p \), such as (7a), cannot be shown to be inconsistent in Hintikka’s system.

\begin{equation}
(7a) \quad \text{It isn’t raining in Chicago, but it may be raining there.}
\end{equation}

What can be shown is that (7a) has the same peculiar property as (24a)—Moore’s famous example—and (24b):

\begin{enumerate}
\item \( \text{The cat is on the mat, but I don’t believe it.} \)
\item \( \text{The cat is on the mat, but I don’t know whether the cat is on the mat.} \)
\end{enumerate}
The latter is formally represented in Hintikka’s system by (25):

\[(25) \quad p & \& \sim K_ap & \& \sim K_a \sim p \]

\[p \text{ and } a \text{ doesn’t know that } p \text{ and } a \text{ doesn’t know that } \sim p \]

Hintikka’s explanation of the absurdity of (24b) runs as follows. When somebody makes a statement, we are normally entitled to expect that he is in the position to know that what he is saying is true. At least we would not expect the speaker to deprive himself of this possibility by the very form of the expression he is using. But the statement in (24b) is just that kind of utterance—Hintikka calls it an “epistemically indefensible” statement. It can be shown in Hintikka’s system that it is impossible for the speaker himself to know that (24b) is true. If (24b) is prefixed with I know that, the corresponding formula in (26) turns out to be inconsistent. (Assume that a refers to the speaker himself):

\[(26) \quad K_a(p & \& \sim K_ap & \& \sim K_a \sim p) \]

The proof of the inconsistency of (26)—not explicitly presented by Hintikka—is given in (27):

\[(27) \quad 1. \quad K_a(p & \& \sim K_ap & \& \sim K_a \sim p) \in W \quad \text{I.P.} \\
2. \quad p & \& \sim K_ap & \& \sim K_a \sim p \in W \quad 1, \ (C.K) \ [\text{see (22)}] \\
3. \quad \sim K_ap \in W \quad 2, \ (C.&K) \ [\text{see (4)}] \\
4. \quad P_a \sim p \in W \quad 3, \ (C.\sim K) \\
5. \quad \sim p \in W^o \quad 4, \ (C.P^o) \\
6. \quad K_a(p & \& \sim K_ap & \& \sim K_a \sim p) \in W^o \quad 5, \ (C.KK^o) \\
7. \quad p & \& \sim K_ap & \& \sim K_a \sim p \in W^o \quad 6, \ (C.K) \\
8. \quad p \in W^o \quad 7, \ (C.&K) \]

Lines 5 and 8 show that, given the initial assumption on line 1, there would have to be a model set that contained both \(p\) and \(\sim p\), which is a violation of the condition \((C.\sim)\) given in (4).

As you can observe by studying the proof in (27), the third conjunct plays no role in it. Thus, the same proof also explains the absurdity of the expression in (28a), which in turn is interchangeable with (28b):

\[(28) \quad a. \quad p & \& \sim K_ap \quad p \text{ but I don’t know that } p \\
\quad b. \quad p & \& P_a \sim p \quad p \text{ but, for all I know, it is possible that } \sim p \]

Now, (28b) is just the kind of utterance that was first introduced in (7). Although Hintikka does not explicitly discuss sentences of this type, it is clear that in his system they have the same status as the
examples in (24). The sentences in (7) are not inconsistent as such, but they are nevertheless epistemically indefensible—that is, self-defeating for anyone to utter, since the truth of the statement can never be known by the speaker.

The crucial condition in Hintikka’s system is \( \text{C.}KK^o \) given in (22d). It is intuitively much less obvious than the other conditions, but without it the desired result would not come about. It is interesting to note that this condition is closely related to the principle that we first presented in (6) and later unsuccessfully tried to formalize in (17). The condition \( \text{C.}KK^o \) requires that the subset of formulas that \( a \) knows in some possible world be included in all of its epistemic alternatives as well. In simple terms, the condition can be rephrased as in (29):

(29) **Whatever is known to be cannot (epistemically) be otherwise.**

Hintikka (1962) presents the following justification for the condition:

If it is consistent of me to say that it is possible, for all that I know, that \( q \) is the case, then it must be possible for \( q \) to turn out to be the case without invalidating any of the my other claims to knowledge; that is, there must not be anything inconsistent about a state of affairs in which \( q \) is true and in which I know what I say I know [p. 17].

The interesting point is that in ordinary language, simple unqualified, nonmodal statements, such as *It isn’t raining in Chicago*, carry with them an implicit claim *I know that it is so*, from which it follows that the speaker may not simultaneously admit that it might not be so without violating the rules of discourse. That is, simple declarative statements ought to be not just true, but epistemically necessary for the one who utters them. If this view is correct, it gives some insight in the problem why ordinary language words like *necessarily, must*, and *have to* most of the time seem to play a very different role in modal logic than their supposed counterpart, the necessity operator.

**THE WORD MUST**

There is a striking difference between the logical necessity operator and words like *must*. Consider the two pairs of expressions in (30) and (31):

(30)  

| a. \( Lp \) | Necessarily \( p \) |
| b. \( p \) |
(31) a. John must have left.
b. John has left.

In any of the standard modal logics, $Lp$ is a stronger expression than $p$. However, there is an inverse relation between the two sentences in (31).

Intuitively, (31a) makes a weaker claim than (31b). In general, one would use (31a) the epistemic must only in circumstances where it is not yet an established fact that John has left. In stating (31a), the speaker indicates that he has no first-hand evidence about John’s departure, and neither has it been reported to him by trustworthy sources. Instead, (31a) seems to say that the truth of John has left in some way logically follows from other facts the speaker knows and some reasonable assumptions that he is willing to entertain. A man who has actually seen John leave or has read about it in the newspaper would not ordinarily assert (31a), since he is in the position to make the stronger claim in (31b).

The relation between (31a) and (31b), therefore, is of entirely different nature than the relation between the two logical formulas in (30). If we were to paraphrase (31a) without using the word must, the most accurate paraphrase would probably be something like the following:

(32) From the things I either know or regard as very probable, it logically follows that John has left (although I cannot report this as an established fact).

This claim is supported by the following observations. Suppose you have just uttered (31a) and are now called upon to justify your statement. It seems to me that a successful defense of (31a) would have to be a list of some circumstantial evidence and some axiomatic principles of reasoning from which (31b) logically follows. For example, one might point out facts like those listed in (33):

(33) a. John never goes anywhere without his hat.
b. John’s hat was on the shelf earlier.
c. John’s hat is not on the shelf now.
d. Nobody else but John could have removed that hat.

The defense of (31a) is successful, provided that the premises given by the speaker are generally accepted by others and provided that the truth of (31b) in fact follows from those premises. Note also that the claim made by (31b) cannot be defended at all by logical reasoning. To defend (31b) one must be able to say something like I saw
him leave or I know that he did. Anything short of that amounts to giving up (31b), it is a retreat to (31a). In trying to defend (31b) by reasoning, one essentially admits having made a stronger claim than one was entitled to make. On the basis of such observations, it seems clear that in statements like (31a) must is not to be interpreted as It is logically necessary that. . . . Just like possible in the earlier examples, it represents a weaker epistemic notion. The relation between the epistemic must and epistemic possible is given in (34):

\[(34) \quad \text{For all I know, it must be that } p \]
\[= \text{For all I know, it is not possible that } \sim p.\]

A similar observation was made by Gottlob Frege (1879, translated in Geach and Black, 1966), who used it as an argument for excluding modal notions from pure logic altogether. In Frege’s words:

What distinguishes the apodeictic from the assertoric judgment is that it indicates the existence of general judgments from which the proposition may be inferred—an indication that is absent in the assertoric judgment. If I term a proposition “necessary”, then I am giving a hint as to my grounds for the judgment. But this does not affect the conceptual content of the judgment; and therefore the apodeictic form of a judgment has not for our purposes any significance [Geach and Black, 1966, p. 4].

Thus, the role of must in (31a) is to indicate that the complement proposition is inferred but not yet known to be true independently. The intuitive feeling that (31a) is a weaker assertion than (31b) is apparently based on some general conversational principle by which indirect knowledge—that is, knowledge based on logical inferences—is valued less highly than “direct” knowledge that involves no reasoning.

It is obvious that what is true about must is also true of other similar words, such as necessarily (when it modifies a sentence), have to, or is bound to. The imperfect correspondence between these words and the logical necessity operator has been commented on by many logicians, some of whom unfortunately regard it as just another defect of ordinary language. The following syllogism is from Gerald J. Massey (1969, p. 183), who presents it as an example of how the word necessarily is sometimes sloppily used in places where it does not properly belong:

\[(35) \quad \text{If there is a sea fight tomorrow, then the Athenians will be victorious.}
\]
\[\text{But the Athenians will not be victorious.}
\]
\[\text{Necessarily, therefore, there will be no sea fight tomorrow.}\]
However, far from being an example of the misuse of *necessarily*, (35) is a perfect example of the ordinary epistemic use of this word. Following Aristotle, some logicians draw a distinction between relative and absolute necessity. Like Aristotle, they say that the conclusion of a syllogism is necessary in relation to its premises. Given a valid argument form and true premises, the conclusion cannot be false. However, at the same time, the conclusion may fail to be necessary in the absolute sense of the term. The relevant passage in Aristotle’s Prior Analytics reads as follows: “One might show by an exposition of terms that the conclusion is not necessary simply, although it is necessary in relation to the premises” [i. 10(30b 32–33), quoted from Kneale and Kneale, 1962, p. 93]. It seems to me that Aristotle’s distinction between relative and absolute necessity is precisely the distinction between epistemic and logical necessity. If I know that the premises of (35) are true, then it is not consistent of me to say that, for all I know, it is possible that the conclusion is false.

However, it has been argued by some logicians (e.g., Kneale and Kneale, 1962, p. 93) that, while it is important to be aware of the ambiguity of *necessarily* in order to avoid logical fallacies, ultimately the distinction can be eliminated by simply reanalyzing every case of relative necessity as a case of absolute necessity.

In both Greek and English, the same words are used to express both absolute and relative necessity and possibility, and this is natural; for every case of relative necessity and possibility can be expressed as a case of absolute necessity and possibility. In syllogistic, for example, if a conclusion is necessary in relation to certain premises, it is absolutely necessary that if the premises are true, the conclusion should be true also. Similarly, if the truth of \( p \) is possible in relation to \( q \), it is absolutely possible that \( p \) and \( q \) should be true. But this use, though natural, may give rise to confusion, because statements of relative necessity and possibility are often made elliptically and may for this reason be misunderstood as statements of absolute necessity and possibility [Kneale and Kneale, 1962, p. 93].

As far as I can see, this reanalysis amounts to saying that, in (35), *necessarily* is predicated, not of the conclusion, but of the syllogism as a whole. That is, it is absolutely necessary that, if the premises are true, then the conclusion is also true.

The same advice has also been given with regard to conditional sentences. For example, Hughes and Cresswell (1968, p. 27) discuss the example in (36):

\[
(36) \quad \text{If it rains throughout December, it is bound to rain on Christmas Day.}
\]
Although the authors at one point correctly interpret (36) to mean that it will rain on Christmas Day follows from it will rain throughout December, they want to identify the phrase it is bound to with the logical necessity operator. They deplore the fact that the structure of the English sentence looks very much like (37b). For them, the logical form of (36) is the formula given in (37a):

\[(37) \quad \begin{align*}
    \text{a. } & L(p \supset q) \\
    \text{b. } & p \supset Lq
\end{align*}\]

As they see it, the phrase it is bound to is properly used only if it expresses logical necessity. Therefore, they suspect that “most frequently what the speaker intends to assert (or at least all he is entitled to assert) is something of the form \(L(p \supset q)\).”

Whether the solution really works for (36) is somewhat doubtful. If we read (36) as if it had the form (37a), the sentence at first appears to be a convincing case of logical necessity. On second thought, however, one realizes that the matter depends on whether Christmas Day simply means “December 25” or “the birthday of Christ.” In the latter case, of course, If it rains throughout December, it will rain on Christmas Day is not logically necessary. Its truth depends on the unstated premise that Christ was born in December, which presumably is a contingent fact. Furthermore, in most cases of this type, changing the scope of the necessity operator does not help at all. It does not convert relative necessities to absolute necessities. Consider the example in (38):

\[(38) \quad \text{If Bill has a diamond ring, he must have stolen it from someone.}\]

It is clear that must in (38) does not express logical necessity, no matter whether it is predicated of the conditional as a whole or of the consequent alone. It is simply a mistake to think that epistemic modalities can be made to collapse into logical modalities by re-arranging the constituent structure of the sentence.

**MODALS IN CONDITIONALS**

Although the change in the scope of the modal in (38) does not result in a logically necessary statement, it is interesting to observe that (38) apparently is equivalent in meaning with (39):

\[(39) \quad \text{It must be that if Bill has a diamond ring, he has stolen it from someone.}\]
Consider also the following pair of examples:

(40)  a. *This year Easter is in April. Necessarily, therefore, if it rains throughout April, it will rain on Easter Sunday.

b. This year Easter is in April. Therefore, if it rains throughout April, it will necessarily rain on Easter Sunday.

It does not seem to make any difference in meaning whether the epistemic necessarily has the conditional as a whole in its scope—as in (40a)—or whether the scope of necessarily is limited to the consequent clause. On the other hand, note that the two formulas in (37), which supposedly correspond to the ordinary language conditionals in (40), are not equivalent in any of the standard modal logics. This is what we would expect, provided that necessarily in (40) is analyzed as an epistemic operator and not as a logical modality. If (40a) is true, then the truth of If it rains throughout April, it will rain on Easter Sunday follows from the truth of This year Easter is in April. But this is just the same as to say that the truth of It will rain on Easter Sunday follows from This year Easter is in April and It will rain throughout April. In other words, whenever a set of premises $P$ jointly entail a conditional “If $A$ then $B$,” then the truth of $B$ logically follows from the set $P \cup \{A\}$, and vice versa. To make this explanation work, one must of course assume that the ordinary language conditional in (40) is not a truth functional operator. A theory of conditionals that seems capable of accounting for these phenomena has recently been proposed by Stalnaker (1968). Following a suggestion by F. P. Ramsey, Stalnaker proposes that a conditional statement, such as If it rains throughout April, it will necessarily rain on Easter Sunday in (40b), is to be evaluated by the following method. You add (hypothetically) the antecedent to your stock of knowledge, and then consider whether the consequent holds. The judgment about the conditional as a whole should be the same as the judgment about the consequent in the hypothetical context. Since the truth of This year Easter is in April is known in (40b), the addition of the antecedent to the stock of knowledge makes it epistemically necessary that it will rain on Easter Sunday. If I know that Easter is in April and that it is going to rain throughout April, then it is not consistent of me to say that it possibly won’t rain on Easter Sunday.

Stalnaker’s theory can also account for the fact that the indicative conditional in (41b) is anomalous in just the same way as (41a):

(41)  a. *It is raining in Chicago but it is possible that it isn’t raining there.
(41)  b. *It is raining in Chicago, but if it isn’t raining there, then the weather forecast was right.

In ordinary language, any indicative conditional normally commits the speaker to the view that, for him, it is epistemically possible for the antecedent to turn out to be true; in Stalnaker’s terms, the addition of the antecedent to this stock of knowledge must not result in inconsistency. Therefore, (41b) is epistemically indefensible for anyone to utter, just as (41a) is. As already observed, a straight non-modal statement like It is raining in Chicago not only commits the speaker to the belief that it is so, he ought to know that it is so. This can be seen by comparing (41b) with the two examples in (42):

(42)  a. I believe that it is raining in Chicago, but if it isn’t raining there, the weather forecast was right.
    b. *I know that it is raining in Chicago, but if it isn’t raining there, the weather forecast was right.

Example (41b) is just as anomalous semantically as (42b) where I know that is explicitly stated. The contrast between (42a) and (42b) shows clearly that, if the logic of ordinary language conditionals is an extension of modal logic, as Stalnaker proposes, the logic of indicative conditionals in particular must refer to epistemic modalities. The antecedent of an indicative conditional ought to be epistemically possible.

Conditionals with possible bring up the same problem that we saw in (40). Although the two formulas in (43) are not equivalent in

4 By normally, I mean to exclude certain ironic uses of conditionals, such as If Harry is a genius, then I am a monkey’s uncle, which consist of an antecedent that the speaker considers false followed by an absurd consequent. The point of such utterances obviously is to make the hearer realize that, since the consequent cannot be true, the conditional as a whole can be true only if the antecedent is false. It is an indirect way of communicating that Harry is not a genius. This special use of conditionals is nicely explained by truth-functional analysis, but it can be accounted for in Stalnaker’s analysis as well by postulating an “absurd” world in which anything is true (Stalnaker and Thomason, 1970). There are similar cases, which, however, seem to call for a somewhat different explanation. Namely, one may temporarily suspend one’s own convictions for the sake of a polite argument and pretend that some statement that one is about to refute is possible. The refutation consists of showing that the acceptance of the statement leads to a conclusion that even the opposition ought to recognize as intolerable. For example, one might try to persuade a lexicalist of the error of his ways by saying: “If your theory is correct, then Her hit I is a grammatical sentence.” In arguing for the truth-functional analysis of conditionals, Grice (1968) presents a couple examples that also require the speaker to act as if he didn’t know what he knows. I do not think that they count as real counterexamples to my claim that an indicative conditional commits the speaker to the view that the antecedent is epistemically possible.
standard modal logic, similar pairs in ordinary language apparently are synonymous:

(43)  
   a. \( M(p \supset q) \)  
   b. \( p \supset Mq \)

(44)  
   a. It is possible that, if he can get it cheap, Sam will smoke pot.  
   b. If he can get it cheap, it is possible that Sam will smoke pot.

The two examples in (44) have been discussed by George Lakoff (1970). On the basis of their synonymy, Lakoff argued that the two sentences are to be derived from the same underlying structure, such as (45), where the if clause comes at the end.

(45) It is possible that Sam will smoke pot, if he can get it cheap.

In Lakoff’s analysis, there would be a transformation that could optionally move the if clause either to the beginning of the complement clause—yielding (44a)—or all the way to the beginning of the main clause. Lakoff does not consider the problem of whether the antecedent is originally inside or outside the scope of possible. He seems to take the first alternative for granted, apparently for no better reason than the fact that, were it not so, the analysis would incorrectly predict an ambiguity in (44b). Otherwise, (44b) could have two interpretations, one corresponding to (43b), where the if clause has been preposed but not raised, the other corresponding to (43a) with both preposing and raising. On the other hand, (44a) could have only one possible source. Since there is no such ambiguity in (44b), Lakoff must somehow rule out as inadmissible all structures of the type (43b), where the antecedent is outside the scope of the modal.

Once it is understood that possible in (44) refers to an epistemic notion and does not correspond to the possibility operator, one may begin to look for a more satisfactory solution. Given Stalnaker’s semantics for conditionals, there apparently is no need to try to explain the synonymy of the two examples by deriving them from the same underlying structure. Both sentences are true just in case Sam will smoke pot is possible in a situation where we know that he can get it cheap, in addition to what else we know in the actual world.\(^5\) Thus,
there is no reason to assume that the scope of the modal in the underlying structure of (44b) is different from its scope in the surface structure.

LOGICAL VERSUS EPISTEMIC MODALITIES

Thus far I have presented examples where ordinary language modals like possible and must represent epistemic modalities and should not be confused with their counterparts in modal logic. If it is true that modals in ordinary language are usually interpreted epistemically, the question arises as to how one expresses in ordinary language that something is logically possible or logically necessary. This brings us back to the distinction between (7a) and (8a’):

(7a) *It isn’t raining in Chicago, but it may be raining there.
(8a’) It isn’t raining in Chicago, but it could be.

In trying to explain why (7a) was contradictory while (8a’) was not, we suggested that (8a’) meant that it could be raining in Chicago if something were not the way it is. In other words, given that it is not raining in Chicago, it is not epistemically possible that it is raining there; nevertheless, it may logically be possible. It is the logical sense of possibility that is conveyed by (8a’). Consider now (46):

(46) It isn’t raining in Chicago, and it couldn’t be.

The word couldn’t in (46) carries the sense of logical impossibility. What (46) says is that even if things weren’t the way we know they are now, it still wouldn’t be raining in Chicago. Thus, something is impossible in the strong logical sense if it would not be possible even if I knew differently. The mood of the verb can be used to distinguish between epistemic and logical possibility.

However, in many cases there is a genuine ambiguity between epistemic and logical modality, just as Aristotle observed. For example, (47) can be taken in two ways:

(47) It isn’t necessarily raining in Chicago.

In the epistemic sense (47) means “For all I know, it doesn’t follow that it is raining in Chicago.” The logical sense of (47) could be paraphrased as “Even if it is raining in Chicago, it could as well be otherwise.” Some similar cases are disambiguated by the presence of polarity items. For example, (48) has only the epistemic reading:

(48) *He ain’t necessarily that bright.
It will be interesting to investigate how the distinction between logical and epistemic modalities is manifested in ordinary language. At this stage, we do not yet know much about it.

REFERENCES