V | Conventions of Language

1. Possible Languages

A verbal signaling language £ has been called a language, and rightly so. But it is a rudimentary language, for at least the following reasons.

There is only a closed, finite set of sentences of £. Truth conditions can be given sentence by sentence. There is no way to create a new sentence, with its truth condition, out of old parts.

There is no such thing as idle conversation. Sentences of £ are reserved for use during some particular activity, for the sole purpose of carrying on that activity successfully. If £ is used by a convention, therefore, the convention is sustained by an interest in coordination in the short run. Anyone's failure to conform if the rest conform leads directly to an audience's response which is undesired given the state of affairs that holds.

Users of £ have little free choice. In the indicative case, a communicator who has observed a certain state of affairs and wishes to be truthful in £ must utter a certain sentence. He has no choice whether to speak or be silent; no choice what to talk about; no choice even how to phrase his message.

For any indicative sentence of £, we have stipulated that the communicator is in a position to tell whether it is true in £ on a particular occasion. There is no place for indicative sentences that express personal opinions or tentative hypotheses.

An indicative sentence is true in £ in an instance of a certain situation if a certain state of affairs holds in that instance. Indicative sentences never express general facts—only facts about the occasion of utterance of the sentence. (For the reasons listed so far, indicative sentences of £ are inadequate for carrying on speculation, deliberation, or argument.)

The audience is in a position to make true any imperative sentence of £ by a response within the situation in which it is uttered. There are no imperative sentences that convey general advice or standing orders.

The audience has an interest in making true any imperative sentence of £ uttered by a communicator who expects it to be made true. There is no need to decide whether the communicator has the knowledge and wisdom to give sound advice, or the authority to issue orders, or the power and will to enforce demands, and so on.

There are not enough moods. Though it is easy to multiply moods beyond necessity, there is no plausible way to get along with only our two. Perhaps the most urgent additional moods are: interrogative; commissive, for promises and threats; permissive, for explicit witholding of imperatives.[1] Some of these may be reducible to indicatives and imperatives, but all such reductions are problematic at best.

There is no ambiguity or indexicality in £. Each sentence has a single fixed mood and a single fixed truth condition, the same on every possible occasion of utterance within £.

Verbal signaling languages may be deficient in still more ways. But we have seen enough. Let us turn to a larger class of possible languages. I hope it will be large enough to contain languages like Welsh or English or Esperanto. If not, I hope it will at least contain large central fragments of these languages.

In this section, we shall consider what a *possible* language is, in abstraction from any users it might happen to have, and in abstraction

[1] See R. M. Hare, "Some Alleged Differences between Imperatives and Indicatives," *Mind,* 76 (1967), pp. 309–326, for an account of permissives as explicitly withheld imperatives. Hare suggests that we might distinguish a mood analogously related to the indicative mood: "a way of volubly and loquaciously *not* making a certain statement" (p. 321).

even from the question of how it *would* be used if it had users. Later we shall consider what it is for a population to use a language—in other words, what makes a *possible* language someone's *actual* language.

By "possible language" I simply mean an entity of the sort I am about to specify. I do not intend to say that every such entity is a serious candidate to be the actual language of any human population. There will be possible languages whose use could serve no human purpose, possible languages so clumsy that they never could be adopted, and even possible languages that men are psychologically or physiologically unable to use.

By a language, I mean an *interpreted* language. So what I call a language is what many logicians would call a language plus an interpretation for it. For me there cannot be two interpretations of the same language; but there can be two languages with the same sentences.

We shall start from the possible languages we have seen: verbal signaling languages. Such a language, we recall, is a function $\mathfrak{L}$ which assigns to every verbal expression in some finite set—every sentence of $\mathfrak{L}$—an interpretation consisting of a mood μ and a truth condition τ. The mood μ is a code number 0 for indicatives, 1 for imperatives. The truth condition τ is a set of possible instances of the signaling problem to which $\mathfrak{L}$ applies: namely, those instances in which the sentence is true either in the indicative sense (if μ is 0) or in the imperative sense (if μ is 1). Now we must amend this description step by step for the sake of generality.

In the first place, we have seen that there should be more than two moods. Let us allow a possible language to have any finite number of moods, since we do not know an upper bound on the number that will be needed. But since there probably is an upper bound, and probably not a very high one, we will get many uninteresting so-called possible languages. That does not matter; what would matter would be to leave out interesting ones. These new moods may simply be further code numbers beyond 0 and 1: say, 2 for interrogatives, 3 for commisives, 4 for permissives, and so on, in some order for whatever others are needed. If, as we hope, some of these moods are reducible to indicatives and imperatives, we can simply leave some code numbers unused.

What difference is there between the moods, to justify us in calling them indicative, imperative, and the like? None, for the time being. As long as we abstract from the use of a possible language, moods are nothing but numbers occurring in interpretations. But later, when we see what it takes to make a possible language be an actual language of a population, different code numbers will play different roles in the conventional regularities whereby the language is used. Only then will the names we have attached to the code numbers be justified. We have already seen how this happens for a verbal signaling language: when the language is used by convention, it is the communicator who tries to make his sentences true if μ is 0, the audience that tries to make the communicator's sentences true if μ is 1.

In the second place, we can no longer take truth conditions to be sets of possible instances of the situation to which the language applies; there no longer *is* any special situation to which the language applies. Instead, we can take a truth condition to be a set of possible worlds: the set of those worlds in which, as we say, the truth condition holds. If $\mathfrak{L}$ assigns an interpretation $\langle \mu, \tau \rangle$ to a sentence σ, then τ is the set of worlds in which σ is true in the sense appropriate to μ.

In the third place, $\mathfrak{L}$ may contain indexical sentences whose truth conditions depend on the utterer, on his intended audience, or on the time and place of utterance. $\mathfrak{L}$ may contain anaphoric sentences whose truth conditions depend on the context of previous conversation or intended subsequent conversation. (That is, sentences like "Then he took off his coat" or "The aforesaid party refused to pay.") $\mathfrak{L}$ may contain sentences whose truth conditions depend on the surroundings of their utterance: "Close the door" or "There is salt on the table." So $\mathfrak{L}$ must assign interpretations not to sentences themselves, but to sentences on possible occasions of their utterance. A possible occasion of utterance of a sentence σ may perhaps be identified with a pair of a possible world and a spatiotemporal location therein, such that σ is uttered at that location in that world. Given the world and location, presumably all further information we need about the context will be forthcoming. We will have uniquely identified the utterer, his intended audience, the previous conversation, the surroundings, and so on. (In those bizarre possible worlds in which more than one utterance of a sentence can occur at one location, more information about the context would have to be built into the entity we take to be the occasion of utterance. But there is no need for all occasions of utterance of sentences to be entities of the same type.)

In the fourth place, we cannot assume that sentences of $\mathcal{L}$ are unambiguous. So $\mathcal{L}$ must assign not an interpretation but a set of interpretations. When the sentence is unambiguous on an occasion of utterance, the set will contain just one interpretation; when the sentence is ambiguous, the set will contain finitely many interpretations. It is probably convenient to allow a third case: the set might contain *no* interpretations. That is one way to treat sentences like "The door is open" on an occasion of its utterance when no door either is present or has been mentioned; and anomalous sentences like "The kettle is dead," "Every girl sang himself a horse," or "Quadruplicity drinks procrastination" on all occasions of their utterance. (An alternative would be to treat the anomalous sentences as nonsentences, at the cost of complicating the grammar of the language; but that would not take care of "The door is open," since we do not want sentencehood itself to be relative to occasions of utterance. Another alternative would be to assign interpretations with empty truth conditions, as we would to self-contradictory sentences. That might be best; but I prefer to leave the question open.) We might like to stipulate that all the interpretations assigned to a sentence on an occasion of its utterance should have the same mood; but we should not. Consider "Shut the door" on an occasion of its utterance when a prominent open door is making everyone chilly, and when the utterer has also been telling a story and has just been asked "What did he do next?"

By now we have accumulated a large change in the character of the function $\mathcal{L}$. We now have a function whose arguments are pairs of a sentence and a possible occasion of its utterance, and whose values are finite, possibly empty, sets of interpretations. An interpretation is still a pair of a mood and a truth condition; but a mood may now be any natural number (though it will be a small one if £ is an interesting language) and a truth condition is a set of possible worlds.

Originally it would have been reasonable to say that the meaning of a sentence was given by the single interpretation assigned to it by £, independently of its occasions of utterance. But we can no longer say that the meaning—or even *a* meaning—of a sentence is given by any one of its interpretations. This will not even do for a sentence that is unambiguous on every occasion of its utterance. For no indexical or anaphoric features of the sentence will show up in the interpretation it receives on any one occasion. What *does* give the meaning of the sentence—insofar as that can be done without considering the meanings of its parts—is the function whereby its set of interpretations depends on features of occasions of utterance.

Finally, we can no longer stipulate that the set of sentences of the language—the domain of £—is finite. Any interesting language has infinitely many sentences. Admittedly, at most a finite number of sentences of any language will happen actually to get uttered. At most a finite number *could* get uttered in any possible world in which human limitations remain as they are. But these finite sets are of no interest, since we have no hope of finding out which sentences they contain. It is a commonplace that any user of a language has the same competence regarding sentences that never happen to be uttered as he has regarding sentences he meets every day. Whenever this is so, his language should be taken to include all the sentences he *could* use and understand, had he enough patience, time, and memory. Unless we allow languages with infinitely many sentences, we will have to impose arbitrary and unjustifiable upper limits on the length of sentences.

2. Grammars

Not just any arbitrary infinite set of verbal expressions will do as the set of sentences of an interesting language. No language adequate to the purposes of its users can be finite; but any language usable by finite human beings must be the next best thing: finitely specifiable. It must have a finite grammar, so that all its sentences, with their interpretations, can be specified by reference to finitely many elementary constituents and finitely many operations for building larger constituents from smaller ones.

A good deal of recent effort in linguistic theory has been devoted to finding a suitable normal form for grammars.[2] The plan is to cut down the class of possible languages by cutting down the class of possible grammars, until the only possible languages left are the ones that are serious candidates for human use. In practice, this is done by restricting the normal form of grammars as far as can be done without leaving any actual language grammarless.

It will do us no harm to have many extra entities counting as possible languages, as well as the ones we really want. So we will not stipulate that our possible languages must have grammars of any specified form. In fact, we need not include possession of *any* grammar as a defining condition for possible languages. We can save ourselves the trouble of trying to say, with adequate precision and generality, what it is to have a grammar. But we must bear in mind that languages without grammars—or without grammars of whatever turns out to be the appropriate normal form—are called possible languages only because we have been too lazy to rule them out.

Let me nevertheless try to say how one sort of grammar for a possible language $\mathfrak{L}$ might work. I distinguish three parts of the grammar, called the lexicon, the generative component, and the representing component.

The *lexicon* is a large finite set of elementary constituents, marked to indicate their grammatical categories. Most of these will be words, or morphemes smaller than words.

The *generative component* is a finite set of *combining operations*. Each of these operates on a given number of constituents of given

[2] See, for instance, Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge, Mass.: MIT Press, 1965), chap. 1.

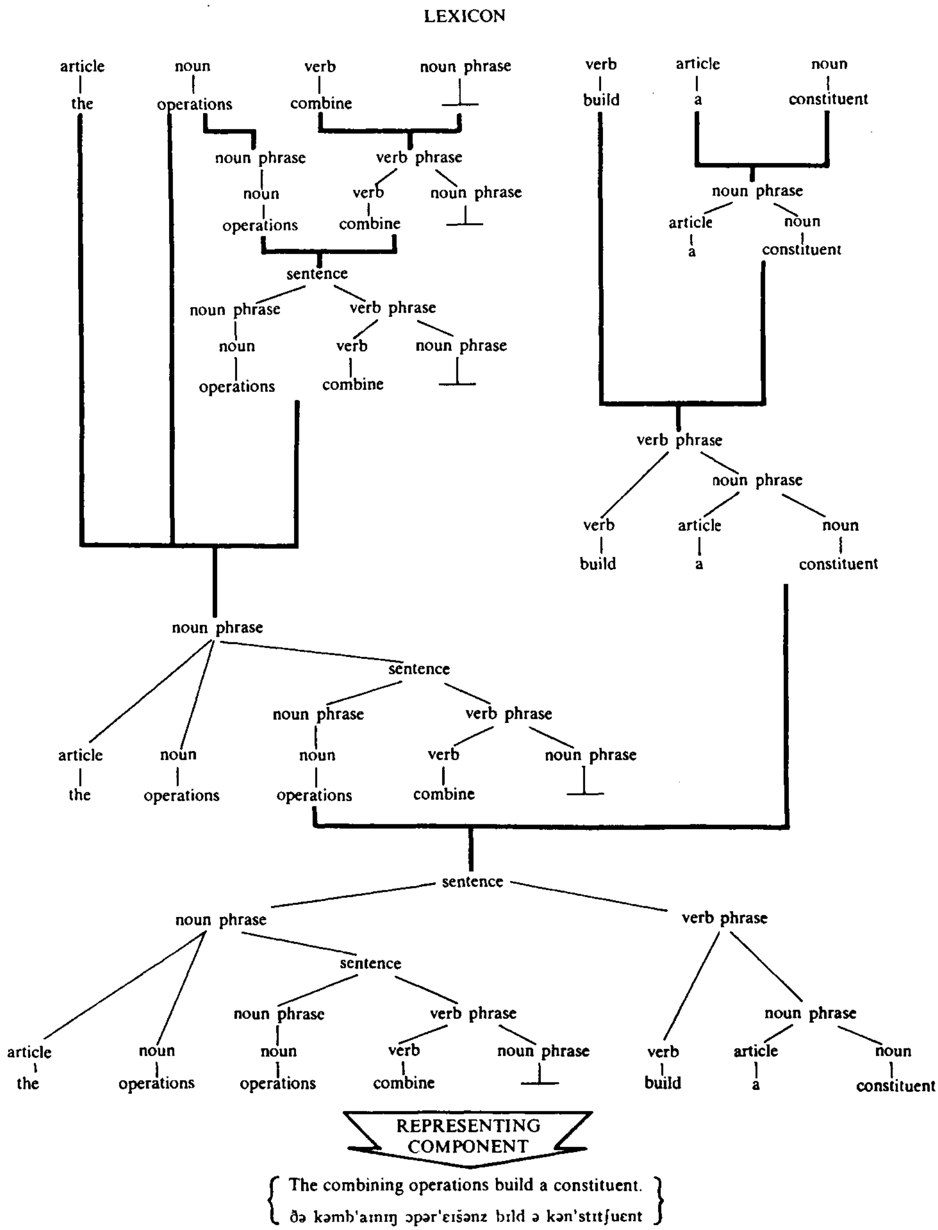
LEXICON
article
the
noun
operations
verb
combine
noun phrase
verb
build
article
a
noun
constituent
noun phrase
verb phrase
sentence
REPRESENTING COMPONENT
The combining operations build a constituent.
ðə kəmb'aɪnɪŋ ɔpər'eɪšənz bɪld ə kən'stɪtʃuent

Figure 46

categories, concatenating them to build a new, larger constituent of a given category. It also provides this new constituent with a marker showing its category and the constituents of which it is built. Starting with the lexical elements, the generative component builds up larger and larger constituents. More precisely: a constituent is any member of the smallest set containing the lexical elements and closed under the combining operations. Thus constituents are strings of lexical elements, carrying a hierarchy of category markers, as shown in Figure 46.

The *representing component* operates on some of the constituents built by the generative component—those of the category *sentence*—to produce verbal expressions. The verbal expressions thus representing sentential constituents are the sentences of $\mathfrak{L}$. In the special case of a *phrase-structure grammar*, the representing component has little work to do. It merely strips off the category markers and replaces the lexical items, in order, by suitable strings of sounds or marks. Grammars for formalized languages—at least, those with simple systems of punctuation—are phrase-structure grammars. In the more general case of a *transformational grammar*, the representing component does much more. Using information contained in the category markers, it may permute parts of the sentential constituent, delete parts, and add new parts, before it produces a verbal expression.

I do not stipulate that there must be a one-to-one correspondence between sentential constituents and the verbal expressions representing them. One sentence might represent several different sentential constituents: syntactic ambiguity. Or one sentential constituent might be represented by several different sentences: one kind of stylistic variation. Or a sentential constituent might fail to be represented; by permitting the representing component to be selective, we can simplify the generative component.

The grammar should give not just the sentences of $\mathfrak{L}$ but also their interpretations on their occasions of utterance. It can do this by (1) assigning interpretations—we shall soon consider what sort of things these are—to the lexical elements, (2) providing, for each combining operation used to build a new constituent ξ out of old ones $\xi_1 \ldots \xi_k$, an accompanying *projection operation* to derive an interpretation for ξ, given an interpretation for each of $\xi_1 \ldots \xi_k$, and (3) passing on the interpretations of sentential constituents to the sentences representing them. As the combining operations build up infinitely many larger and larger constituents, starting with the lexical elements, the corresponding projection operations work in parallel to derive interpretations for those constituents, starting with interpretations of the lexical elements.

A sentence may have more than one interpretation passed on to it; this can happen in either or both of two ways. The sentence may be syntactically ambiguous, representing—and receiving interpretations from—more than one sentential constituent. An example is Chomsky's "John was frightened by the new methods," which is ambiguous although it contains no ambiguous word. Or the sentence may be ambiguous because it represents one sentential constituent that already has several interpretations. For a lexical element may be ambiguous. When it is built into a larger constituent, the different interpretations of the lexical element will in general yield different intepretations of the larger constituent. And so on, up to an ambiguous sentential constituent, represented by an ambiguous sentence. An example is "Owen is going to the bank."

We need not assume that when a new constituent ξ is built out of old ones $\xi_1 \ldots \xi_k$, *every* possible combination of interpretations of $\xi_1 \ldots \xi_k$ yields an interpretation for ξ. If that were so, ambiguity would run wild. The projection operations may be selective, working only on input combinations of interpretations which satisfy certain restrictions. But in that case, it could happen that ξ received no interpretation at all, although it was properly built up from $\xi_1 \ldots \xi_k$, all of which did have interpretations. For no combination of interpretations of $\xi_1 \ldots \xi_k$ might be an acceptable input for the projection operation. If ξ is built in turn into another constituent, that too will normally have no interpretation; for the projection operation will not be given even one complete combination of interpretations to work on. And so on, up to sentential constituents without interpretations. If such an anomalous sentential constituent is represented by a sentence, it will have no interpretation to pass on to the sentence. That is why it is possible for sentences themselves to have no interpretations. (I do not say that any good grammar for any familiar language would yield sentences without interpretations; but, as I said earlier, it seems advisable to leave the possibility open.)

A lexical element may be indexical, receiving different interpretations on different possible occasions of its utterance. (Let us say that an occasion of utterance of a constituent is any occasion of utterance of a sentence representing a sentential constituent containing it.) This dependence on occasion is passed along through projection operations to larger and larger constituents, and finally to sentences.

My sketch of the nature of a grammar for $\mathcal{L}$ has been designed to have enough generality to cover two special cases: (1) the sets of formation and valuation rules used by logicians to specify the formalized languages they study, and (2) transformational grammars for natural languages, of the form recently proposed by Chomsky, Jerrold Katz, and others.[3] There are earlier proposals by Chomsky and his associates which would have allowed grammars not fitting my description; but no convincing case has been made that this original generality is needed. In any case, it does not matter whether I have given an adequate definition of a grammar, provided I have shown roughly what one would look like.

I have subscribed to Katz's account of the way in which a grammar derives interpretations for a sentence by starting with interpretations for lexical elements, using projection operations to derive interpretations for larger and larger constituents, and finally handing over interpretations from sentential constituents to the sentences repre-

[3] Chomsky, *Aspects of the Theory of Syntax;* Jerrold Katz, *Philosophy of Language* (New York: Harper and Row, 1966), chap. 4. My constituents, since they carry hierarchies of category markers, are the same as Chomsky's underlying phrase markers, or subtrees thereof.

senting them. But I have not endorsed Katz's account of the nature of these interpretations; that is a separate question. Katz takes them to be expressions built out of symbols called "semantic markers" which represent "conceptual elements in the structure of a sense."[4] I find this account unsatisfactory, since it leads to a semantic theory that leaves out such central semantic notions as truth and reference.

Then what is an interpretation for a constituent? We have already decided one case. An interpretation for a sentence, and hence for a sentential constituent, is a pair of a mood and a truth condition: a code number and a set of possible worlds. The mood is something peculiar to sentences; but the truth condition suggests a general strategy for providing constituents with appropriate interpretations.

Referential semantics in the tradition of Tarski and Carnap provides constituents with extensions appropriate to their categories: truth values for sentences, denotations for names, sets for one-place predicates, sets of *n*-tuples for *n*-place predicates, and so on. Given extensions for lexical elements, appropriate extensions for larger and larger constituents are derivable by projection operations (valuation rules). Interpretations had better not be mere extensions, of course, since the extension depends both on the interpretation and on accidental facts about the actual world; for instance, constituents that ought to have different interpretations turn out to be accidentally coextensive. Nevertheless, referential semantics looks like a near miss.

A truth condition specifies truth values for a sentence; but in all possible worlds, not just in whichever world happens to be actual. We can interpret a constituent of any category on the same principle, by giving it an extension (appropriate to its category) in every possible world. The idea is Carnap's; it has recently been applied to the semantics of formalized languages with intensional operators, in work by several philosophers in the tradition of Tarski and Carnap.[5]

[4] Katz, *Philosophy of Language*, pp. 155-156.

[5] See Rudolf Carnap, *Meaning and Necessity*, 2nd ed. (Chicago: University of Chicago Press, 1956), pp. 181-182; Jaakko Hintikka, "Modality as Referential Multi-

For instance, an interpretation for a name should give the thing named (if any) in every possible world. It can be taken as a function from possible worlds to things therein. An interpretation for a one-place predicate should give the things of which that predicate is true in any given world; it can be taken as a function from worlds to sets of things therein (or, if nothing inhabits more than one world, perhaps just as a single set containing things from various worlds). An interpretation for an *n*-place predicate can be taken as a function from worlds to sets of *n*-tuples of things therein. It is possible to provide interpretations even for constituents that resist treatment within referential semantics confined to the actual world. An interpretation for a modal operator, for instance, might be taken as a function from possible worlds to sets of truth conditions—that is, sets of sets of possible worlds.

It is this sort of interpretation—an assignment of extension in every possible world—which, I suggest, should be attached to constituents by the grammar, in order to build up to sentence interpretations of the kind we want. (Some sort of special provision would have to be made in the grammar for attaching moods.) Obviously such interpretations—like extensions in a single world—are capable of being given relative to features of occasions of utterance; for that reason, a meaning for a constituent is not any one interpretation, but rather the

plicity," *Eripainos Ajatus,* 20 (1957), pp. 49-64; Saul Kripke, "Semantical Considerations on Modal Logic," *Acta Philosophica Fennica,* 16 (1963), pp. 83-94; David Kaplan, *Foundations of Intensional Logic* (Ann Arbor: University Microfilms, 1964); Richard Montague, "Pragmatics," *Contemporary Philosophy—La Philosophie Contemporaine,* ed. Raymond Klibansky (Florence: La Nuova Italie Editrice, 1968); Montague, "On the Nature of Certain Philosophical Entities," *The Monist,* 53 (1969); Dana Scott, "Advice on Modal Logic," presented at the Free Logic Colloquium held at the University of California at Irvine, May 1968. Montague and Scott propose a unified treatment of intension and indexicality in which extensions are assigned relative to *points of reference:* combinations of a possible world and several relevant features of context—a time, place, speaker, audience, etc. Montague classifies his work as pragmatics because of this relativity to context; it does not deal with the sort of pragmatic considerations that determine which possible language is used by a given population.

function whereby its interpretation (or its set of alternative interpretations) depends on its occasions of utterance.

3. Semantics in a Possible Language

This completes our examination of possible languages. We can define certain semantic properties of sentences relative to possible languages in general, as we did relative to verbal signaling languages. But to allow for indexicality and ambiguity, these will in general be four-place relations: between a sentence σ of $\mathfrak{L}$, the language $\mathfrak{L}$, a possible occasion o of utterance of σ, and an interpretation $\langle \mu, \tau \rangle$ assigned by $\mathfrak{L}$ to σ on o. If σ is *eternal* in $\mathfrak{L}$—assigned by $\mathfrak{L}$ the same set of interpretations on every possible occasion of its utterance—we can omit mention of the occasion o. If σ is *unambiguous* in $\mathfrak{L}$ on o—assigned a single interpretation by $\mathfrak{L}$ on o—we can omit mention of the interpretation $\langle \mu, \tau \rangle$.

> σ is *indicative* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and μ is 0. (Likewise for the four other moods named.)
>
> σ is *true* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in—that is, contains—the possible world w in which the possible occasion o of utterance of σ is located.
>
> σ is *false* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ does not hold in the possible world w in which o is located.

In speaking of the truth and falsehood of eternal sentences of $\mathfrak{L}$, either we will mention a possible world w or we will be speaking of the actual world.

> σ is *true* in $\mathfrak{L}$ in world w under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to σ on every possible occasion of its utterance a set of interpreta-

tions containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in the possible world w. (Likewise for falsehood.)

> σ is *true* in $\mathfrak{L}$ under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to σ on every possible occasion of its utterance a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in the actual world. (Likewise for falsehood.)

In the simplest case, we can ascribe truth in $\mathfrak{L}$ to an unambiguous eternal sentence of $\mathfrak{L}$. In the absence of contrary stipulation, this is to be taken as truth in the actual world.

> σ is *true* in $\mathfrak{L}$ if and only if $\mathfrak{L}$ assigns to σ on every possible occasion of its utterance a single fixed interpretation $\langle \mu, \tau \rangle$ and the truth condition τ holds in the actual world. (Likewise for falsehood.)

Truth conditions assigned by $\mathfrak{L}$ may be universal, holding in—that is, containing—every possible world; or they may be empty, holding in no possible world; or they may be in between, holding in some but not others. Accordingly, we may call sentences analytic, contradictory, or synthetic in $\mathfrak{L}$, on occasions and under interpretations.

> σ is *analytic* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in every possible world.

> σ is *contradictory* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in no possible world.

> σ is *synthetic* in $\mathfrak{L}$ on o under $\langle \mu, \tau \rangle$ if and only if $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$ and the truth condition τ holds in some possible worlds but not in others.

Again, we can simplify the definienda in speaking of sentences that are eternal in $\mathfrak{L}$, unambiguous in $\mathfrak{L}$ on o, or both.

> σ is *analytic* in $\mathfrak{L}$ if and only if $\mathfrak{L}$ assigns to σ on every possible occasion of its utterance a single fixed interpretation $\langle \mu, \tau \rangle$ and the truth condition τ holds in every possible world. (Likewise for contradiction and syntheticity.)

Note that an unambiguous indexical sentence may be true in $\mathfrak{L}$ on every possible occasion of its utterance without being analytic in $\mathfrak{L}$ on any occasion. Sentences reputed to have this status in English include "I am here now," "I exist," "I am awake," "I am uttering something," and the like.[6] Such a sentence, on any possible occasion of its utterance, is assigned a truth condition that holds in the world in which the occasion of utterance is located, but does not hold in some other possible world. It is reassuring to find that we have not confused this status with analyticity, as would be all too easy to do.

There is an opposite possibility that is not reassuring and comes as a surprise. An unambiguous indexical sentence may be analytic on one possible occasion of its utterance, but false on another. Take the English sentence "It is a perfect square." Take an occasion of its utterance on which the only entity under discussion in the previous conversation was the number 49, referred to by the numeral "49." Then I take it that the truth condition assigned to our sentence on that occasion will hold in every possible world in which 49 is a perfect square—that is, in every possible world. So it is the universal truth condition. The sentence is analytic in English on that occasion. But take another occasion of its utterance in which the entity under discussion was the number 48, referred to by the numeral "48." On

[6] All these familiar examples are subject to objections. (1) "Here," "now," and the present tense might refer, in some cases, not to the place and time of the occasion of utterance but rather to the place and time of the intended occasion of hearing. (2) A ghost who claimed to exist might do so falsely, if there is a sense in which to exist—for a human or ex-human—is to be alive. (3) A man talking in his sleep might succeed in performing the action of uttering something. I mention these objections not to uphold or attack them, but only to say that I am interested in the *status* that has been ascribed to my examples—not in whether they are really good examples of that status.

that occasion, our sentence is not analytic in English; it is false, and indeed contradictory, in English.

I find nothing wrong with this. Recall that most discussion of analyticity has ignored analytic indexical sentences; we ought to have no firm expectations about them. Moreover, although our sentence is analytic in English on certain occasions, it is not analytic in English *simpliciter*, according to the definition given; but some other indexical sentences like "Yesterday is past" are not only analytic in English on occasions of their utterance, but also analytic in English *simpliciter*. Any possible occasion of utterance of the sentence "Yesterday is past" is located on some day d. The truth condition assigned to the sentence on that occasion holds in every possible world in which day $d-1$ precedes day d; this is the universal truth condition, no matter what d is. So our sentence is assigned a single fixed interpretation on every possible occasion of its utterance, and this interpretation contains the universal truth condition.

Customarily, a sentence of any mood may be called contradictory; but only an indicative sentence may be called analytic. I have ignored this pointless restriction: analyticity—like contradiction, truth, and falsehood—depends on the truth condition assigned by $\mathfrak{L}$, without regard to the accompanying mood. An analytic imperative is, "Wear a hat or else don't!" An analytic commissive is "I promise to remain unmarried so long as I am a bachelor." An analytic permissive is "You may respire whenever you breathe." (For reasons to be considered later, I doubt that there is any such thing as an analytic question.)

This completes my quick, but highly general, account of possible languages and of the semantic properties of sentences therein. We turn now to the question: if $\mathfrak{L}$ is a possible language, and P is a population of agents, what relation must hold between $\mathfrak{L}$ and P in order to make it the case that $\mathfrak{L}$ is an actual language of P? The obvious answer is: the members of P must *use* the language $\mathfrak{L}$. But that answer is neither informative nor clear. They must use $\mathfrak{L}$ in a certain way. If everyone in P used $\mathfrak{L}$ by telling lies in $\mathfrak{L}$ or by singing operas in $\mathfrak{L}$ (without understanding the words), they would be using

£ but not in the right way; £ would not be their language. Only when we already know what makes a language be an actual language of a population can we say what kind of use we had in mind.

The next answer is: the members of *P* must give to the sentences of £ the interpretations (on occasions of utterance) which are assigned to them by £. But this is another uninformative answer. What sort of action is it to give an interpretation to a sentence? Not something anyone can do just by putting his mind to it. I can't say "It's cold here" and mean "It's warm here"[7]—at least, not without a little help from my friends.

4. Conventions of Truthfulness

We are trying to find out what the members of a population *P* must do in order to make it the case that a certain possible language £ is their actual language—that they *use* £, in the sense we are after. It is surely something they do in conformity to a convention: something everyone in *P* does because he expects his conversational partners in *P* to do it too, and because a common interest in communicating leads him to want to do his part if they do theirs. This much we know, just because we know that it matters little (in the long run) which language we use, so long as we all use the same one.

I shall contend that the convention of language whereby the members of a population *P* use a given possible language £ may best be described as a convention of truthfulness in £.[8] That is the conclusion we already reached for the special case in which £ was a verbal

[7]Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 1958), sec. 510.

[8]I owe this proposal to Erik Stenius, "Mood and Language Game," *Synthese,* 17 (1967), pp. 254–274. He proposes that language use is governed by rules giving truth conditions for sentence-radicals (sentences minus their indicators of mood) and by rules prescribing the appropriate sort of truthfulness for each mood. He considers three moods: indicatives, imperatives, and yes-no interrogatives. I have adapted his proposal by building truth conditions into the identification of possible languages and by taking his rules of truthfulness as conventions of truthfulness. Stenius in turn acknowledges a debt to Dr. H. Johansen of Copenhagen for "the idea that speaking the truth can, in a sense, be regarded as a semantic rule."

signaling language. In that case, we deduced the existence of a convention of truthfulness in £ from (1) our previous account of signaling conventions, and (2) our application of the concept of truth to conventional signaling. Now we shall consider how less restricted languages also might be used in conformity to conventions of truthfulness.

Let us postpone consideration of the nonindicative moods, of indexicality, and of ambiguity. We then have a language £ that assigns to each of its sentences a single interpretation with the indicative mood, the same on every possible occasion of utterance of that sentence. To be truthful in £ is to try not to utter any sentence of £ that is not true in £; but it is not only that. One cannot be truthful in £ just by never uttering sentences of £. (I am not now being truthful in Welsh.) To be truthful in £, in a more positive sense, is to participate in verbal exchanges by occasionally uttering sentences of £, while trying not to utter any that are not true in £. One can, of course, be silent at any particular point in a conversation without ceasing to be truthful in £; but hold your tongue too long—or at too high a cost to your social purposes—and you gradually turn from a truthful user of £ (who happens not to want to say anything) to a nonuser of £.

We did not have to distinguish between positive and negative truthfulness with respect to verbal signaling languages, because the signaling convention left no choice whether to speak or to be silent. To conform, one must utter whichever sentence of the language is true. But it would be absurd to claim that our conventions of language prohibit silence at any point in an ordinary conversation. And yet anyone who *always* chooses silence cannot be said to conform to our conventions of language.

If our simple language £ is the actual language of a population P, we will certainly find a regularity R of truthfulness in £. Members of P will exchange utterances of sentences of £, and they will almost always try to avoid uttering sentences not true in £. (Sometimes they will lie; but we can tolerate exceptions to a conventional regularity. Sometimes they will be mistaken; but the regularity I have in mind is that of *trying* to be truthful in £, not of succeeding. Conventions are regularities in choice of action.) Moreover, it will be common knowledge in *P* that members of *P* almost always conform to *R*. How about the other conditions for the conventionality of *R*?

Any member of *P* will (almost always) want to conform to *R* if other members of *P* do; they also will want him to do so; and these preferences will be common knowledge in *P*. We can safely assume that £ is a useful language, since otherwise it could not be the actual language of *P*. Then uniform truthfulness in £ permits successful communication, which answers to all sorts of interests of the members of *P*. Consider any ordinary conversation among members of *P*. Someone wants to get the others to share some of his beliefs. It does not matter why he wants them to have those beliefs—he might want them to act on those beliefs, for selfish or altruistic reasons; he might be trying to educate them; he might just be passing the time of day. He can accomplish his purpose by uttering sentences he takes to be true in £—that is, by the proper kind of conformity to *R*. And the reasons he can accomplish his purpose by conforming to *R* is that others have conformed to *R* in the past: namely, those who shaped the habits and expectations of his present audience. Because of the prevailing regularity of truthfulness in £, his audience has become accustomed to truthfulness in £. They are habitually truthful in £. They habitually act and form beliefs on the assumption that others are truthful in £. Being practiced in £, they do so quickly and easily. In this way, members of *P* normally have reason to be truthful in £, conditionally upon recent widespread truthfulness in £ by members of *P*.

On this account, the coordination achieved by uniform truthfulness in £ is diffuse, one-sided coordination among communicators who, at different times, communicate with the same audience. Each communicator wants to be truthful in £ because that is what past communicators have led the audience to expect. A member of the audience, as such, is not constrained by convention. He merely listens and perhaps forms beliefs, in the knowledge that the communicator is probably being truthful in $\mathfrak{L}$. Only when he takes his turn as communicator does he himself act in conformity to the convention of truthfulness in $\mathfrak{L}$.

The case of a verbal signaling language with indicative signals was described differently. The convention of truthfulness is two-sided: the communicator tries to utter a sentence true in the language, and the audience acts as seems best on the assumption that he has succeeded in doing so. The coordination sought is coordination between communicator and audience.

I think we are right to treat the two cases differently. Verbal signaling is carried out with some definite end in view, and it is common knowledge what that end is. But this is not the case in general. I do not tell you only what you need to know right now in order to serve our common purposes. Often there is nothing in particular that the audience should do if the communicator has told the truth. The audience should form a belief, perhaps, but that is normally not a voluntary action and hence not an action in conformity to convention. Even if the audience should act, the action may not answer to an interest common to the communicator and the audience, may not be one the communicator intended to produce, and so on; and hence it may not be plausibly describable as constrained by convention. No doubt there is a continuous spectrum from verbal signaling to idle chat, and two-sided and one-sided coordination may be mixed in various proportions. But generality is served by concentrating on the one-sided coordination among communicators which is present in all conventional indicative communication, not on the occasional two-sided coordination between a communicator and his audience.

Finally, a convention must have an alternative. If R is a convention of truthfulness in $\mathfrak{L}$, its alternatives are regularities of truthfulness in suitable alternative possible languages. For several reasons, not every other possible language is a suitable alternative. Remember that many bizarre entities are still being called possible languages because we did not go to the trouble of ruling them out. We will not be able even to define truthfulness with respect to an arbitrary possible language, but only with respect to one that confines itself to the first few moods and does not have too much ambiguity. Even so, we will have grammarless languages, trivial languages, unpronouncable languages, unlearnable languages, languages whose shortest sentences take hours to utter, languages in which no sentence says anything anyone would ever want to say, and so on. An alternative language to £ must be a possible language for which truthfulness is definable, and one that is sufficiently convenient and useful to make it a serious candidate for human use. But it must not be too close to £, since it must almost always be impossible to act in conformity both to a convention and to its alternative. If £ and its alternatives shared some sentences, and assigned those sentences overlapping truth conditions, it might become possible to be truthful in £ and in its alternative at the same time. Still we need have no fear that the convention of truthfulness in £ will lack an alternative. There are all the actual languages of different populations, and plenty of other possible languages that would serve as well as they do.

We have now completed our survey of the defining conditions that must be met if there is to be a convention in *P* to be truthful in the simple language £; and we have found that, on the supposition that £ is the actual language of *P*, these conditions would be met. There would prevail in *P* a convention of truthfulness in £. This convention would be sustained by a certain sort of interest on the part of members of *P*: an interest in communication, in being able to control one another's beliefs and actions, to some extent, by means of sounds and marks.

One is inclined to deny that there can be a convention of truthfulness in £, on the grounds that truthfulness is a moral obligation. This objection rests in part on confusion between truthfulness in a given possible language £ and truthfulness in the language of one's population, whichever possible language that may be. It is the former, not the latter, which is conventional. But both are obligatory, according to our common moral opinions. I am obligated to be truthful in £, on condition that my fellows are truthful in £; I am obligated to be truthful in whichever language they use, on condition that they do use some language or other. I grant that these are moral obligations, but I deny that they prevent truthfulness in £ from being conventional.

Why might they seem to? I suppose because conventions, as they have been defined, are sustained by preference. To conform to a convention of truthfulness in £ is to be truthful in £ because it answers to one's own preferences to do so, given that others do. But virtuous people feel bound by an obligation to be truthful in £ if others are, even if that goes against their preferences.

This much is true: one who is truthful in £ against his own preferences cannot then be acting in conformity to a convention. But such cases are exceptional. In the world as we know it—and as it must be, if use of language is to persist among sinful men—almost everyone almost always has reason to get others to share his beliefs, and therefore has reason to conform to conventions of truthfulness. Thus in the normal case, one can both be fulfilling a moral obligation and be acting according to one's preferences. Whenever that is so, we can have a case of conformity to convention. Only some exceptional cases need to be described otherwise.

A convention of truthfulness in £ is a social contract as well as a convention. Not only does each prefer truthfulness in £ by all to truthfulness in £ by all but himself. Still more does each prefer uniform truthfulness in £ to Babel, the state of nature. So each ought to recognize an obligation of fair play to reciprocate the benefits he has derived from others' truthfulness in £, by being truthful in £ himself. In the exceptional cases, this obligation will be his only reason to be truthful in £. In the normal cases, it will be present but redundant, since he will also have sufficient self-interested reasons.

A different objection can be made to my contention that conventions of language are conventions of truthfulness in a given possible language £. Recall what £ is: a certain function whose arguments are pairs of a sentence (a finite sequence of types of sounds or of marks) and a possible occasion of its utterance (a pair of a possible world and a spatiotemporal location therein) and whose values are sets of interpretations (pairs of a code number and a set of possible worlds). Now it is incredible that any ordinary user of £ has a concept of any such complicated entity. So how can he be party to any convention regarding £? How can we have expectations and preferences regarding truthfulness in £?

My answer is that he can have them *in sensu diviso*; to do that, he does not need to share our complicated concept of a possible language. He does not even need to share our concepts of function, sequence, set, pair, and possible world. He can be a nominalist philosopher who disbelieves in all of these things or a peasant who has never heard of any of them. All he has to do is to come up with the right expectations and preferences regarding particular instances of what we—but not he—would call truthfulness in £. And these particular instances are pieces of perfectly commonplace human thought and action. The only explicit use of esoteric concepts is ours, when we summarize an enormous and varied body of belief-governed action by classifying it under an invented description: truthfulness in £.

It is logically possible, then, to conform to a convention of truthfulness in £ without having any general concept of truthfulness in £. This is not to say that it is *psychologically* possible to do so without having something closely analogous to that concept; and I mean to say no such thing. The user of £ is a finite being with very limited experience; yet somehow he has acquired an enormous, and enormously varied, repertory of propensities to action, expectation, and preference in a wide variety of situations. That is what it takes for him to be—as we but not he would say—habitually truthful in £ and accustomed to expect truthfulness in £ on the part of others. I can imagine no explanation of this competence which does not posit some sort of unconscious counterpart of a general concept of truthfulness in £.

Consider a bicycle rider. Any credible theory of his competence will ascribe to him unconscious mental processes closely analogous to the use of general physical concepts, in the knowledge of general physical laws. But unless our bicycle rider happens also to be a physicist, it would be wrong to say that he had those concepts or knew those laws, although his expectations regarding a wide variety of particular situations would work according to those laws.

This completes my account of conventions of truthfulness in simple languages without nonindicative moods, indexicality, or ambiguity. Still postponing indexicality and ambiguity, let us turn to truthfulness in languages with imperatives, interrogatives, commissives, and permissives.

Recalling our discussion of imperatives in verbal signaling languages, we can think of imperative truthfulness roughly as obedience. That is, it is up to the audience to make imperative sentences of £ true in £, by trying to act in such a way that the truth conditions assigned to them by £ will hold.

As a first approximation, we might consider this statement of a convention among members of *P* of truthfulness in £, insofar as it applies to imperatives: if any member of *P* thinks he has been the intended audience of an utterance by another member of *P* of an imperative sentence σ of £, then he tries to make σ true in £.

This is not a very good approximation. We certainly have no convention to make true just any imperative. There are bad advice, unauthorized orders, exorbitant requests, unenforceable demands, and so on. All of these are imperatives that we can ignore without violating a convention. It is only when the communicator stands in a certain *relation of authority* to his intended audience that the audience will violate convention if they fail to try to make the communicator's imperatives true. At least in many cases, and perhaps in all, the appropriate relation of authority is as follows: the communicator and his audience have—and it is common knowledge between them that they have—a common interest in making it possible for the communicator to control the audience's actions, at least within a certain range, by uttering verbal expressions.

This relation holds between the communicator and the audience in any signaling problem with verbal signals, by our definition of a verbal signaling problem. It holds whenever the communicator is advising the audience, and it is common knowledge that he is in a position to give good advice and wants to do so. It holds whenever the communicator gives the audience an order in the course of performing some cooperative task, if it is common knowledge that this is what he is doing. It holds whenever the communicator makes a request that is not exorbitant and it is common knowledge that the audience wants to please him.

But this relation of authority does not seem to hold between a communicator and an audience when the communicator's imperative is a demand enforced by a threat. The opposite holds: it is common knowledge that the audience would be better off if the communicator were *not* able to control their actions. Perhaps an audience giving in to a demand should not be described as conforming to a convention of truthfulness, although we can concede that demanding by means of imperatives is possible because of the existence of conventions of truthfulness governing those imperatives in other situations, situations in which the communicator does stand in a relation of authority to the audience. Or perhaps we should say that demands are best analyzed not as imperatives but as commissives—threats—even when they are grammatically like imperatives.

We have arrived at the following statement of a convention of truthfulness in $\mathfrak{L}$, insofar as it applies to imperatives. If a member x of P thinks he has been the intended audience of an utterance by another member y of P of an imperative sentence σ of $\mathfrak{L}$, and if y and x have—and it is common knowledge between them that they have—a common interest in making it possible for y to control x's actions within a certain range by uttering verbal expressions, and if there is some action in that range whereby x could try to make σ true in $\mathfrak{L}$; then x does try to make σ true in $\mathfrak{L}$.

I said that a convention of truthfulness in $\mathfrak{L}$ constrained the communicator of an indicative, but not the audience; likewise I am now representing the convention as constraining the audience of an imperative, but not the communicator. But in this case I see no good reason not to include the communicator: his part is to utter imperatives as seems best, assuming truthfulness in $\mathcal{L}$ on the part of his audience. That is, when he stands in the relation of authority to an audience, he utters those imperatives that he wants the audience to make true in $\mathcal{L}$ and that the audience can make true in $\mathcal{L}$ by actions within the acceptable range of control. On the other hand, it does not seem clearly necessary to include the communicator's part in a statement of the convention. We can just as well represent him as doing what seems best, in the knowledge that the convention exists and that his audience will probably conform to it. As in the purely indicative case, we do not need to concentrate on the short-term, two-sided coordination between a communicator and his audience. There is also a long-run, diffuse coordination among all those who converse with any given person. For whoever converses with him both influences his linguistic habits and accomplishes various conversational purposes—or fails to accomplish them—by virtue of the linguistic habits he already has.

Turning to interrogatives, we can save ourselves some work. Instead of treating interrogatives as a new mood, we can treat them as a species of imperative. (Therefore the code number 2 reserved at the beginning of this chapter for the interrogative mood will not occur in interpretations assigned by the languages we are considering.) Take the question "Is dinner ready?"; compare it with the imperative "Tell me whether dinner is ready!" or, more explicitly, "Say 'yes' if dinner is ready, 'no' if not!" Or take the question "What is your name?"; compare it with the imperative "Say your name!" Or take the question "Tell me, who first reached Greenland?"; compare it with the imperative "Tell me who first reached Greenland!" I can find no important differences in any of these pairs. So I propose that questions are imperatives—imperatives with a distinctive subject matter, marked by a distinctive grammatical form. Like any imperative, a question is made true in $\mathcal{L}$ by its audience's performance of the commanded action: the action we call giving a true answer to the question. A true question is a truly answered question.[9] Truthfulness in £ with respect to questions consists in trying to give true answers, at least when the questioner stands to one in a relation of authority. This is simply a special case of truthfulness in £ with respect to imperatives. At the same time, it is a special case of truthfulness in £ with respect to indicatives; the answers—at least when they are complete sentences—are indicatives, as well as being actions that make imperatives true.

I said in passing that I doubted there were analytic questions. If questions are imperatives, analytic questions are imperatives that cannot fail to be made true; for sentences of any mood are analytic in £ if and only if they are assigned the universal truth condition. So they are questions that cannot fail to be answered truly. But any question properly so-called has contrary alternative answers. Moreover, one can fail to answer truly by keeping silent and so failing to answer at all.

The commissive mood of promises and threats is akin to the indicative. It is up to the utterer to be truthful in £, by trying to see to it that the truth condition of his commissive holds. To make a promise or threat true in £ is to fullfill it, to do what one has said one was going to do. My indicative "I shall return" and my commissive "I will return" are both true in English if and only if I subsequently return; and to be truthful in English, I must try to make sure that I utter either sentence only if, later, I do return. That is, I must try to act in such a way that there is a correspondence between my words now and my deeds later. But there are two different times at which I may try to make sure of this correspondence: now, when I choose my words, or later, when I choose my deeds.

In the case of the indicative "I shall return," to be truthful is to

[9] Here I disagree with Stenius. He proposes, in "Mood and Language Game," that the sentence-radical of a yes-no question be considered true if and only if "yes" is the correct answer. Thus he prescribes this sort of truthfulness for questions: "Answer the question by 'yes' or 'no,' according as its sentence-radical is true or false" (p. 273). Certainly that is the most natural notion of truth for yes-no questions; but unlike mine, it cannot be extended to other kinds of questions.

try *before* I utter the sentence to make my words correspond to my deeds. I do this by trying to foresee or decide whether I will return, and by refraining from uttering "I shall return" if I foresee or decide that I will not. It is irrelevant to my truthfulness that I do or do not go on trying to make sure of the correspondence after my utterance. (It *is* relevant to whether my sentence is true; but to be truthful is to *try*—not necessarily succeeding—to make sentences true.) Rash self-prediction is untruthful; unforeseen violation of a self-prediction is not.

In the case of the commissive "I will return," to be truthful is to try *after* I utter the sentence to make my words correspond to my deeds. I do this by remembering my words and accordingly trying to return. It is irrelevant to my truthfulness—though relevant to whether my sentence is true—that I did or did not try before my utterance to make my words correspond to my deeds. Unforeseen failure to try to keep a promise is untruthful; rash promising is not.

If members of a population P are truthful in $\mathcal{L}$, and σ is an indicative sentence of $\mathcal{L}$, they will try before uttering σ to judge whether the truth condition assigned to σ by $\mathcal{L}$ holds. If σ is a commissive sentence of $\mathcal{L}$, they will try after uttering σ to act in such a way as to bring it about that the truth condition assigned to σ by $\mathcal{L}$ holds. Suppose they have a convention of truthfulness in $\mathcal{L}$. Then in doing either of these things, they are conforming to convention. They *may* do other things too: try to make indicatives true which have already been uttered, try not to utter commissives they foresee will not be made true. But these actions are not covered by the convention of truthfulness in $\mathcal{L}$. Among ourselves, they are covered by no convention at all.

A previous discomfort is apt to recur. The whole point of promising—or threatening, as strategists know—is to bind oneself to do something whether or not it turns out, at the time, to answer to one's preferences (so that others' expectations about one's action may be firmer than their expectations about one's preferences). But an action in normal conformity to convention is, by definition, an action that answers to one's preferences. So how can keeping a promise be done in conformity to any convention? It cannot—when it really does go against preference. But when preference agrees with promissory obligation, there is no reason why keeping a promise should not be done in conformity to a convention. The latter is the normal case, and it must be in any population with a convention of truthfulness in a language containing commissives. That should not surprise us: we normally keep promises because we do not want to disappoint others' legitimate expectations; because good will come of it, given what others have done in the expectation that the promise will be kept; because we do not want to incur retaliation, destroy our reputations for keeping promises, or undermine confidence in promises generally. We normally do prefer to keep our promises. And this is a conditional preference, of the kind required if promise keeping is to be action in conformity to convention. If others did not generally try to make true in $\mathcal{L}$ the commissive sentences of $\mathcal{L}$ which they have uttered, no one would have any of the reasons I mentioned for preferring to try to make true in $\mathcal{L}$ the commissive sentences of $\mathcal{L}$ that he has uttered. Any promise has the power to bind an honorable man against his preferences. But most promises are never called upon to exercise that power.

Permissive sentences are neutralizers of imperatives and commissives. If I permit someone to do something (or leave something undone), I ordinarily cancel some command I have given him or some promise he has given me. He can take advantage of my permission by acting in a way that would otherwise have been untruthful, hence contrary to our convention as it applies to my original command or his original promise. Because of my permission, his acting in that way is not untruthful. The cancellation may be complete or partial: having ordered you (or received your promise) to keep off my grass, I may later tell you that henceforth you may come on the grass whenever you like, or I may tell you that you may come on the grass just this once.

I shall not describe a new kind of truthfulness in $\mathcal{L}$, truthfulness with respect to permissives. Rather, permissives will enter the statement of a convention of truthfulness in £ by way of an elaboration of my account of truthfulness with respect to imperatives and commissives. This is not to say that the permissive mood itself will be eliminated. Permissive sentences of £, with interpretations containing the code number 4, will remain. But our statement of a convention of truthfulness in £ will contain no part devoted only to permissives. Indeed, nothing new need be added to our statement of the convention. We need only make an observation about the nature of the truth conditions that may be assigned to imperatives and commissives in a language containing permissives.

The truth condition assigned by £ to a permissive sentence σ gives us the state of affairs which is permitted to hold when σ is uttered: the state of affairs in which, as we say, someone takes advantage of the permission he has been given by an utterance of σ. Thus we know what it is for an uttered permissive to be made true in £.

The simplest sort of truth condition for an imperative or a commissive is the set of just those possible worlds in which a certain person performs a certain action. We can call this a *positive* truth condition. But if an imperative or commissive is conditional—"Take an umbrella if it is cloudy!" or "I promise to help if you need me"—there is also a *negative* truth condition containing just those worlds in which the antecedent state of affairs does not occur. The more antecedents there are, the more different negative truth conditions. The whole truth condition of an imperative or commissive is the union of all its positive and negative truth conditions. In the case of an imperative or a commissive that can be neutralized by a permission, we have another kind of negative truth condition containing just those worlds in which neutralization has occurred. Thus my plan is to take any imperative or commissive that can be neutralized as if it were explicitly conditional upon—*inter alia*— the absence of any neutralizing permission.

More precisely, suppose we have an imperative or commissive whose positive truth condition concerns the action of a certain person

x. This imperative or commissive may also have a negative truth condition that can be described as follows: the set of all possible worlds in which any permissive sentence σ of $\mathfrak{L}$ is uttered to x by anyone in a certain position, and σ is subsequently made true in $\mathfrak{L}$ by x. (Neutralization occurs if the truth condition of σ falls at least partly outside the positive truth condition of the original imperative or commissive; otherwise we get a world lying both within our negative truth condition and within the original positive truth condition.) This negative truth condition will be included in the union of positive and negative truth conditions which is the whole truth condition of the original imperative or commissive.

We need not make any special provision for negative truth conditions of this kind, but only notice that they are possible. And if some imperatives or commissives of $\mathfrak{L}$—not necessarily all—have such truth conditions, then the use of suitable permissives of $\mathfrak{L}$ is already covered by a convention of truthfulness in $\mathfrak{L}$ as it applies to imperatives and commissives.

One may object that there are situations in which permission is required but the requirement was not created by any previous imperative or promise. What situations these are will depend on the institutions of a particular population. We should not expect them to be covered by conventions purely of language. Perhaps they are not covered by any conventions. But we can think of these as by-products of the use of permissions as neutralizers of imperatives and commissives. We can say that the population has a rule or convention to act *as if* certain imperatives or commissives had been uttered: imperatives or commissives with truth conditions allowing for neturalization.

This is as far as we will go in examining truthfulness applied to the nonindicative moods. Doubtless there are more moods to be examined, but I have nothing to say about them and I do not believe they are very important to our understanding of language.

We can easily extend our consideration to a possible language $\mathfrak{L}$ with unambiguous indexical sentences, assigned different interpretations on different possible occasions of utterance. In fact, there is a shortage of eternal nonindicatives in English, so we have already had to use indexical sentences to illustrate various points. Let us take the opportunity to recapitulate the statement of a convention of truthfulness in £, but this time providing throughout for the dependence of both moods and truth conditions upon occasions of utterance. Let *P* be a population in which the convention holds; let *x* be (almost) any member of *P*.

> If σ is a sentence of £ which would be indicative in £ on an occasion *o* of its utterance by *x* to an audience in *P*, then *x* tries to make sure that he utters σ on *o* only if σ would be true in £ on *o*.
>
> If σ is a sentence of £ which was imperative in £ on an occasion *o* of its utterance to *x* by a member *y* of *P*, and if *y* and *x* have—and it is common knowledge between them that they have—a common interest in making it possible for *y* to control *x*'s actions within a certain range by uttering verbal expressions, and if there is some action in that range whereby *x* could try to make σ true in £ on *o*, then *x* tries to act in such a way that σ was true in £ on *o*.
>
> If σ is a sentence of £ which was commissive in £ on an occasion *o* of its utterance by *x* to an audience in *P*, then *x* tries to act in such a way that σ was true in £ on *o*.

As planned, questions and permissives have not been mentioned separately—questions because they are included among the imperatives, permissives because they enter into the truth conditions for some imperatives and commissives.

I have not yet said what truthfulness is with respect to ambiguous sentences of £. Perhaps we do not need to face the question. It might do to say merely that £ is an actual language of a population *P* only if there is a conventional regularity in *P* of truthfulness with respect to those sentences of £ that are unambiguous in £ on their occasions of utterance. This convention might be a consequence of a more general convention of truthfulness in £ with respect to all sentences of £, ambiguous or not; but we could give an account of the more limited convention without knowing what truthfulness in £ is with respect to ambiguous sentences.

Alternatively, we could try to describe a minimal standard of truthfulness in £ with respect to ambiguous sentences. Take a sentence σ which is assigned multiple interpretations by £ on an occasion o of its utterance. One can be minimally truthful in £ with respect to σ on o by taking any one of those interpretations and doing whatever one would have to do to be truthful in £ if that interpretation were the only one assigned to σ on o by £. A trickster is being truthful in this minimal way if, knowing that Owen is going to the shore of the river, he says, "Owen is going to the bank" during a conversation about Owen's lack of cash.

To describe a higher standard of truthfulness, we would have to mention our actual ways of resolving ambiguity in conversational practice. These depend, I suppose, on our opinions about each other's conversational purposes. For instance, we can ignore an interpretation assigned by £ to a sentence σ on an occasion o if that interpretation, but not some other one, is conversationally pointless: if it is common knowledge among the parties to the conversation on occasion o that if that were the only interpretation assigned by £ to σ on o, utterance of σ on o could in no way serve any conversational purpose of the utterer. We can ignore an interpretation of a sentence under which it is common knowledge that the sentence is blatantly false. We can ignore an interpretation of a sentence if it is common knowledge that neither the utterer nor any other party to the conversation notices that the sentence has that interpretation. There are only the crudest of our methods of resolving ambiguity. Yet I hesitate to propose that even these are part of the content of our convention of truthfulness in a language. It seems better to think of them as resulting from our exercise of common sense in the presence of a convention of language. In any case, we cannot build our methods of resolving ambiguity into a definition of truthfulness in £ until we understand better what our methods of resolving ambuity are.

Let us be content, therefore, to take truthfulness in a possible language £ as truthfulness in £ with respect to unambiguous sentences of the moods discussed, and perhaps also minimal truthfulness in £ with respect to ambiguous sentences of those moods.

So far, I have been content to claim that, as a matter of fact, language users are party to conventions of truthfulness. To test my claim, we must simply draw on our knowledge about what goes on when ordinary people use a language in an ordinary way. We can imagine what it would be like for a possible language £ to be an actual language of a population *P* more or less like ourselves, in circumstances more or less like ours. When we imagine this, I maintain, we find that on my definitions there prevails in *P* a convention to be truthful in £—a convention sustained by an interest of the members of *P* in communication, in being able to control one another's beliefs and actions to an extent by producing sounds and marks. We considered this claim first for verbal signaling languages; next for languages simplified by removal of nonindicatives, indexicality, and ambiguity; and finally for languages with these complications restored.

If it is true that conventions of truthfulness are a feature of normal language use as we know it, and if—as I suppose—they are an important feature thereof, then it might be reasonable to appeal to them in defining what it is for a possible language £ to be an actual language of a population *P*. That is, we might adopt the definition:

> £ is an *actual language* of *P* if and only if there prevails in *P* a convention of truthfulness in £, sustained by an interest in communication.

According to our previous discussion of verbal signaling, this definition would be adequate in the special case of verbal signaling languages; I conjecture that it is also adequate in general. That is, I conjecture that it agrees with ordinary usage in clear cases, and draws a convenient line among unclear cases. The test of the definition would be to see if we can think of clear cases in which it disagrees with our inclinations to affirm or deny that £ is an actual language of *P*; and I have not been able to think of any.

It is possible to argue by counterexamples that the definition does not give a necessary condition for £ to be an actual language of P. We can invent many bizarre cases in which we have some inclination to say that £ is an actual language of P although there is no convention in P of truthfulness in £. We can imagine a population of inveterate liars, or of people who suspect each other of being inveterate liars, or of people who use their language only to tell tall tales, or of creatures of instinct who are unable to use any language other than the one that is built into them. But none of these bizarre counterexamples is convincing, for once we appreciate how peculiar they are—how different from language use as we know it—we will not want to classify them as *clear* cases under ordinary usage. And if they are unclear cases, we are free to settle them in whatever way we find convenient. We can happily admit, of course, that they are cases in which a language is, in an extended sense, an actual language of a population; this is simply to say that they bear important resemblances to cases in which the condition given in the definition is satisfied.

5. Semantics in a Population

To the extent that we have given adequate necessary and sufficient conditions for a language to be an actual language of a population, we are in a position to define some semantic relations of verbal expressions to populations of language users. These will depend in the obvious way on the corresponding semantic relations of verbal expressions to possible languages, by way of the relation we have been examining between languages and populations. In general, we will again have four-place relations: this time, among a verbal expression σ, a population P, a possible occasion o of utterance of σ, and an interpretation $\langle \mu, \tau \rangle$.

> σ *receives* from P on o the interpretation $\langle \mu, \tau \rangle$ if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$.

If σ is *eternal* in P—receives the same set of interpretations from P on every possible occasion of its utterance—we can omit mention of the occasion o. If σ is *unambiguous* in P on o—receives a single interpretation from P on o—we can omit mention of the interpretation $\langle \mu, \tau \rangle$.

> σ is *indicative* in P on o under $\langle \mu, \tau \rangle$ if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$, and μ is 0. (Likewise for the other named moods.)
>
> σ is *true* in P on o under $\langle \mu, \tau \rangle$ if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$, and the truth condition τ holds in—contains—the possible world w in which the possible occasion o of utterance of σ is located. (Likewise for falsehood.)
>
> σ is *analytic* in P on o under $\langle \mu, \tau \rangle$ if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to $\langle \sigma, o \rangle$ a set of interpretations containing $\langle \mu, \tau \rangle$, and the truth condition τ holds in every possible world. (Likewise for contradiction and syntheticity.)

Simplifying for the case in which σ is both eternal and unambiguous in P, we get definitions like these:

> σ is *true* in P in world w if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to σ on every possible occasion of its utterance a single fixed interpretation $\langle \mu, \tau \rangle$ whose truth condition τ holds in the possible world w.
>
> σ is *true* in P if and only if there exists a possible language £ such that £ is an actual language of the population P, and such that £ assigns to σ on every possible occasion of its utterance

a single fixed interpretation $\langle \mu, \tau \rangle$ whose truth condition τ holds in the actual world.

σ is *analytic* in P if and only if there exists a possible language $\mathfrak{L}$ such that $\mathfrak{L}$ is an actual language of the population P, and such that $\mathfrak{L}$ assigns to σ on every possible occasion of its utterance a single fixed interpretation $\langle \mu, \tau \rangle$ whose truth condition τ holds in every possible world.

We can simplify the definienda in a different way, retaining mention of the occasion o of utterance of σ, but not mentioning the population P. For by examining o we can identify a communicator and his intended audience, and look for an actual language of a population to which they belong. Thus we get semantic relations between a verbal expression σ and an occasion o of its utterance (and, if we need to provide for ambiguity, also an interpretation). Neither a population nor a language need be mentioned. I will illustrate the simplified definienda without provision for ambiguity.

σ is *true* on o if and only if there exist a possible language $\mathfrak{L}$ and a population P such that the communicator and intended audience on o belong to P, $\mathfrak{L}$ is an actual language of P, and $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a single interpretation $\langle \mu, \tau \rangle$ whose truth condition τ holds in the possible world w in which o is located.

σ is *analytic* on o if and only if there exist a possible language $\mathfrak{L}$ and a population P such that the communicator and intended audience on o belong to P, $\mathfrak{L}$ is an actual language of P, and $\mathfrak{L}$ assigns to $\langle \sigma, o \rangle$ a single interpretation $\langle \mu, \tau \rangle$ whose truth condition τ holds in every possible world.

So much for the interpretations given by a population to those verbal expressions that are the sentences of its language. How about the nonsentential constituents in a grammar for that language? These too carry interpretations. (Let us overlook indexicality, ambiguity, and anomaly, and pretend that they carry single fixed interpretations.)

The grammar assigns interpretations to all its constituents, either directly or by means of its projection operations. But we should like to say that a word, for instance, is given an interpretation by a population of language users. Welshmen give the word "gwyn" a certain interpretation—perhaps it is the function that assigns to every possible world the set of white things therein. In saying this, we have mentioned neither a language nor a grammar. We want a three-place relation: the constituent ξ receives from the population P the interpretation ι.

How can we define this relation? Presumably by mentioning a language $\mathcal{L}$ and a grammar Γ to link the population P with the assigment of ι to ξ. Γ assigns ι to ξ; $\mathcal{L}$ is given by Γ; and P has a convention of truthfulness in $\mathcal{L}$. Unfortunately, $\mathcal{L}$ will not have just one grammar. Different grammars for $\mathcal{L}$ will interpret constituents differently; one of them may assign ι to ξ, another may not. Different grammars for $\mathcal{L}$ may even analyze sentences differently into constituents; ξ may be a constituent in one grammar for $\mathcal{L}$, but not in another. These differences between grammars for $\mathcal{L}$ cancel out; the different grammars give the same sentences, with the same interpretations. Given P, we select $\mathcal{L}$ by looking for a convention of truthfulness; but given $\mathcal{L}$, how can we select Γ? Conventions of truthfulness pertain to whole sentences and leave the interpretations of parts of sentences undetermined. Perhaps we should look for conventions of some other kind, but I cannot think what the content of such a convention might be. It could not simply be a convention to adopt such-and-such grammar or a convention to bestow such-and-such interpretations upon such-and-such constituents. Conventions are regularities in action, and there is no such action as adopting a grammar or bestowing an interpretation (or if there is, it does not occur in normal use of language).

Some grammars for $\mathcal{L}$ are simple, natural, reasonable, easy, good; others are complicated, artificial, contrived, difficult, bad. Perhaps Γ should be the *best* grammar for $\mathcal{L}$, according to some appropriate method of evaluating alternative grammars. We can define a four-place relation among a constituent ξ, a population P, an interpretation ι, and a method of evaluation M:

> ξ *receives* from P the interpretation ι (according to M) if and only if there exist a possible language $\mathcal{L}$ and a grammar Γ for $\mathcal{L}$, such that $\mathcal{L}$ is an actual language of the population P, and such that ξ is a constituent in Γ and Γ assigns ι to ξ, and such that Γ is the best grammar for $\mathcal{L}$ according to the method of evaluation M.

If we can find a method of evaluation that uniquely deserves a privileged status, we can remove the undesired relativity to M. If not, I fear that the notion we want is indefensible: there is no such thing as *the* interpretation given to ξ by P, but only the various interpretations given to ξ by P according to the various alternative methods of evaluating grammars.

Why might some one method of evaluating grammars deserve a privileged status? Perhaps because it is the one that, as Chomsky has conjectured, enters into the psychological explanation of linguistic competence.[10] Suppose that whenever anyone acquires the ability and propensity to be truthful and expect truthfulness in a language, he does so by forming some sort of unconscious internal representation of a grammar for that language; and suppose that at any stage in a child's acquisition of language, his internally represented grammar is the *best* grammar, according to a certain fixed method M of evaluation, that fits the use of language he has observed around him. Thus it is M, together with observation, that determine which grammar a child will internally represent; and the grammar internally represented by native speakers of a language $\mathcal{L}$ will be the grammar of $\mathcal{L}$ that is best according to M.

This psycholinguistic hypothesis is still speculative. I can only say that if it, or something like it, is true, then there is a clear sense in which a constituent can be said to receive an interpretation from a population. If not, we must look elsewhere for a privileged method

[10] *Aspects of the Theory of Syntax,* chap. 1.

of evaluation, or we must give up. To give up would be to accept Quine's thesis of the inscrutability of reference: no matter how much we know about the population *P*, we have no objectively determinate way of interpreting parts of their discourse shorter than sentences.[11]

I hope we now understand what it would be for a verbal expression σ to be analytic in a population *P*. It remains open whether any verbal expression ever *is* analytic in any population. Analyticity as described so far might be a perfectly intelligible status which happens not to be occupied. Similarly, we may know what it would be for a possible language $\mathcal{L}$ to be the actual language of a population; but we do not know that this ever occurs. And I strongly suspect that it does not.

It is often said that analyticity is not sharp. (To say this is quite different from saying that analyticity, sharp or unsharp, is unintelligible.) In any population, say ourselves, most sentences are clearly synthetic, a few trifling ones are pretty clearly analytic, and everything interesting in philosophy and the sciences seems to be somewhere in between. For instance, there seems to be analyticity somewhere in the fundamental principles of dynamics. But where? Conservation of momentum? Action equals reaction? Force equals mass times acceleration? Somewhere else? We cannot tell. Each seems, somehow, partially analytic. But how is partial analyticity possible?

Not primarily because our conventions of language are conventions to less than the highest degree, although they are. It does not seem as if a higher degree of conventionality—fewer cases of untruthfulness, firmer confidence in the truthfulness of others, fewer exceptions of various kinds—would be likely to make analyticity any sharper. And not primarily because it is unclear exactly what possible worlds there are, although it is unclear. (For instance, I have no idea whether any possible world is five-dimensional.) Since analyticity is truth—in the language of a population—in every possible world, uncertainty about the possibility of worlds could certainly be reflected in

[11] *Word and Object,* pp. 68–79; "Ontological Relativity: The Dewey Lectures 1968," *Journal of Philosophy,* 65 (1968), pp. 185–198.

uncertainty about the analyticity of sentences. But that cannot be the whole explanation of unsharp analyticity. Sometimes we cannot tell whether a sentence is analytic—say, one of the principles of dynamics—because we cannot tell whether it is true in our language in some hypothetical world that clearly *is* possible.

I think we should conclude that a convention of truthfulness in a single possible language is a limiting case—never reached—of something else: a convention of truthfulness in whichever language we choose of a tight cluster of very similar possible languages. The languages of the cluster have exactly the same sentences and give them corresponding sets of interpretations; but sometimes there are slight differences in corresponding truth conditions. These differences rarely affect worlds close enough to the actual world to be compatible with most of our ordinary beliefs. But as we go to more and more bizarre possible worlds, more and more of our sentences come out true in some languages of our cluster and false in others. So a sentence could be analytic in some languages of our cluster but false (in sufficiently remote worlds) in others. That sentence would be *partially* or *unsharply* analytic among us. Our actual language is like a resonance hybrid of the possible languages that make it up. Analyticity is sharp in any one of them, but they may not agree.

Even if it did not explain the fact of unsharp analyticity, the hypothesis that our conventions of language restrict us to clusters, not to single possible languages, would be plausible on other grounds. The sort of convention I have in mind is this: almost everyone, almost always, is truthful in at least *some* languages of the cluster: but not necessarily the same ones for everyone, or for one person at different times. This is a better convention for us to have than a convention of truthfulness in a single language. Standardization for the sake of communication is a good thing, but not all-important. If the cluster is tight enough, there is not much threat to communication. Usually we are not talking about remote worlds where truth conditions diverge; so usually truthfulness in any language of the cluster is truthfulness in every language of the cluster. When a confusing divergence does turn up, communication can be preserved—assuming a little good will—by looking for sentences whose truth conditions do not diverge in a troublesome way, and by making temporary conventions more restrictive than our permanent convention. In return for tolerating some threat to communication, we get two benefits.

For one thing, the different languages of the cluster may have different virtues and vices, and hence may be differently suited to individual opinions, tastes, and conversational purposes. If everyone can pick from the cluster, incompatible preferences among languages may all be satisfied. Moreover, by not commiting ourselves to a single language, we avoid the risk of committing ourselves to a single language that will turn out to be inconvenient in the light of new discoveries and theories; we allow ourselves some flexibility without change of convention.

But, more important, the less restrictive our permanent convention is, the less experience it takes to identify it, catch on, and begin to be party to it. A child has to extrapolate from a limited sample of language use, some of it even violating convention. Suppose everyone around him were truthful in exactly the same possible language all the time. The child might still have to wait a long time before he had observed enough conversation to allow him to identify the language. For if two languages differ only in the truth values of sentences in remote possible worlds, the difference will show up in very few conversations. Admittedly, the child is helped by his propensity to ignore almost all the extrapolations consistent with his data—otherwise he could not get anywhere. But provided that more than one language in the cluster is a serious candidate—and this I take to be proved by the existence of unsharp analyticity—he needs more information to identify a language in the cluster than to identify the cluster. So I suggest that a convention of truthfulness in a single possible language could not sustain itself. It would be imperfectly learned; having been imperfectly learned, it would frequently be violated; being frequently violated, it would be still more imperfectly learned.