Languages and Language

Thesis

What is a language? Something which assigns meanings to certain strings of types of sounds or of marks. It could therefore be a function, a set of ordered pairs of strings and meanings. The entities in the domain of the function are certain finite sequences of types of vocal sounds, or of types of inscribable marks; if σ is in the domain of a language £, let us call σ a *sentence* of £. The entities in the range of the function are meanings; if σ is a sentence of £, let us call £(σ) the *meaning of σ in* £. What could a meaning of a sentence be? Something which, when combined with factual information about the world — or factual information about *any* possible world — yields a truth-value. It could therefore be a function from worlds to truth-values — or more simply, a set of worlds. We can say that a sentence σ is *true in* a language £ *at* a world w if and only if w belongs to the set of worlds £(σ). We can say that σ is *true in* £ (without mentioning a world) if and only if our actual world belongs to £(σ). We can say that σ is *analytic in* £ if and only if every possible world belongs to £(σ). And so on, in the obvious way.

Antithesis

What is language? A social phenomenon which is part of the natural history of human beings; a sphere of human action, wherein people utter strings of vocal sounds, or inscribe strings of marks, and wherein people respond by thought or action to the sounds or marks which they observe to have been so produced.

This verbal activity is, for the most part, rational. He who produces

AUTHOR'S NOTE: This paper was originally prepared in 1968 and was revised in 1972. The 1968 draft appears in Italian translation as "Lingue e lingua," *Versus*, 4(1973): 2–21.

certain sounds or marks does so for a reason. He knows that someone else, upon hearing his sounds or seeing his marks, is apt to form a certain belief or act in a certain way. He wants, for some reason, to bring about that belief or action. Thus his beliefs and desires give him a reason to produce the sounds or marks, and he does. He who responds to the sounds or marks in a certain way also does so for a reason. He knows how the production of sounds or marks depends upon the producer's state of mind. When he observes the sounds or marks, he is therefore in a position to infer something about the producer's state of mind. He can probably also infer something about the conditions which caused that state of mind. He may merely come to believe these conclusions, or he may act upon them in accordance with his other beliefs and his desires.

Not only do both have reasons for thinking and acting as they do; they know something about each other, so each is in a position to replicate the other's reasons. Each one's replication of the other's reasons forms part of his own reason for thinking and acting as he does; and each is in a position to replicate the other's replication of his own reasons. Therefore the Gricean mechanism[1] operates: X intends to bring about a response on the part of Y by getting Y to recognize that X intends to bring about that response; Y does recognize X's intention, and is thereby given some sort of reason to respond just as X intended him to.

Within any suitable population, various regularities can be found in this rational verbal activity. There are regularities whereby the production of sounds or marks depends upon various aspects of the state of mind of the producer. There are regularities whereby various aspects of responses to sounds or marks depend upon the sounds or marks to which one is responding. Some of these regularities are accidental. Others can be explained, and different ones can be explained in very different ways.

Some of them can be explained as conventions of the population in which they prevail. Conventions are regularities in action, or in action and belief, which are arbitrary but perpetuate themselves because they serve some sort of common interest. Past conformity breeds future conformity because it gives one a reason to go on conforming; but there is some alternative regularity which could have served instead, and

[1] H. P. Grice, "Meaning," *Philosophical Review*, 66(1957):377–388.

would have perpetuated itself in the same way if only it had got started.

More precisely: a regularity R, in action or in action and belief, is a *convention* in a population P if and only if, within P, the following six conditions hold. (Or at least they almost hold. A few exceptions to the "everyone"s can be tolerated.)

(1) Everyone conforms to R.

(2) Everyone believes that the others conform to R.

(3) This belief that the others conform to R gives everyone a good and decisive reason to conform to R himself. His reason may be that, in particular, those of the others he is now dealing with conform to R; or his reason may be that there is general or widespread conformity, or that there has been, or that there will be. His reason may be a practical reason, if conforming to R is a matter of acting in a certain way; or it may be an epistemic reason, if conforming to R is a matter of believing in a certain way. First case: according to his beliefs, some desired end may be reached by means of some sort of action in conformity to R, provided that the others (all or some of them) also conform to R; therefore he wants to conform to R if they do. Second case: his beliefs, together with the premise that others conform to R, deductively imply or inductively support some conclusion; and in believing this conclusion, he would thereby conform to R. Thus reasons for conforming to a convention by believing something—like reasons for belief in general—are believed premises tending to confirm the truth of the belief in question. Note that I am *not* speaking here of practical reasons for acting so as to somehow produce in oneself a certain desired belief.

(4) There is a general preference for general conformity to R rather than slightly-less-than-general conformity — in particular, rather than conformity by all but any one. (This is not to deny that some state of *widespread* nonconformity to R might be even more preferred.) Thus everyone who believes that at least almost everyone conforms to R will want the others, as well as himself, to conform. This condition serves to distinguish cases of convention, in which there is a predominant coincidence of interest, from cases of deadlocked conflict. In the latter cases, it may be that each is doing the best he can by conforming to R, given that the others do so; but each wishes the others did not conform to R, since he could then gain at their expense.

(5) R is not the only possible regularity meeting the last two condi-

tions. There is at least one alternative R' such that the belief that the others conformed to R' would give everyone a good and decisive practical or epistemic reason to conform to R' likewise; such that there is a general preference for general conformity to R' rather than slightly-less-than-general conformity to R'; and such that there is normally no way of conforming to R and R' both. Thus the alternative R' could have perpetuated itself as a convention instead of R; this condition provides for the characteristic arbitrariness of conventions.

(6) Finally, the various facts listed in conditions (1) to (5) are matters of *common* (or *mutual*) *knowledge*: they are known to everyone, it is known to everyone that they are known to everyone, and so on. The knowledge mentioned here may be merely potential: knowledge that would be available if one bothered to think hard enough. Everyone must potentially know that (1) to (5) hold; potentially know that the others potentially know it; and so on. This condition ensures stability. If anyone tries to replicate another's reasoning, perhaps including the other's replication of his own reasoning, . . . , the result will reinforce rather than subvert his expectation of conformity to R. Perhaps a negative version of (6) would do the job: no one disbelieves that (1) to (5) hold, no one believes that others disbelieve this, and so on.

This definition can be tried out on all manner of regularities which we would be inclined to call conventions. It is a convention to drive on the right. It is a convention to mark poisons with skull and crossbones. It is a convention to dress as we do. It is a convention to train beasts to turn right on "gee" and left on "haw." It is a convention to give goods and services in return for certain pieces of paper or metal. And so on.

The common interests which sustain conventions are as varied as the conventions themselves. Our convention to drive on the right is sustained by our interest in not colliding. Our convention for marking poisons is sustained by our interest in making it easy for everyone to recognize poisons. Our conventions of dress might be sustained by a common aesthetic preference for somewhat uniform dress, or by the low cost of mass-produced clothes, or by a fear on everyone's part that peculiar dress might be thought to manifest a peculiar character, or by a desire on everyone's part not to be too conspicuous, or — most likely — by a mixture of these and many other interests.

It is a platitude — something only a philosopher would dream of denying — that there are conventions of language, although we do not find it easy to say what those conventions are. If we look for the fundamental difference in verbal behavior between members of two linguistic communities, we can be sure of finding something which is arbitrary but perpetuates itself because of a common interest in coordination. In the case of conventions of language, that common interest derives from our common interest in taking advantage of, and in preserving, our ability to control others' beliefs and actions to some extent by means of sounds and marks. That interest in turn derives from many miscellaneous desires we have; to list them, list the ways you would be worse off in Babel.

Synthesis

What have languages to do with language? What is the connection between what I have called *languages*, functions from strings of sounds or of marks to sets of possible worlds, semantic systems discussed in complete abstraction from human affairs, and what I have called *language*, a form of rational, convention-governed human social activity? We know what to *call* this connection we are after: we can say that a given language £ is *used by*, or is a (or the) language *of*, a given population P. We know also that this connection holds by virtue of the conventions of language prevailing in P. Under suitably different conventions, a different language would be used by P. There is some sort of convention whereby P uses £ — but what is it? It is worthless to call it a convention to use £, even if it can correctly be so described, for we want to know what it is to use £.

My proposal[2] is that the convention whereby a population P uses a language £ is a convention of *truthfulness* and *trust* in £. To be truthful in £ is to act in a certain way: to try never to utter any sentences of £ that are not true in £. Thus it is to avoid uttering any sentence of £ unless one believes it to be true in £. To be trusting in £ is to form beliefs in a certain way: to impute truthfulness in £ to others, and thus to tend to respond to another's utterance of any sentence of £ by coming to believe that the uttered sentence is true in £.

[2] This proposal is adapted from the theory given in Erik Stenius, "Mood and Language-Game," *Synthese*, 17(1967):254–274.

Suppose that a certain language £ is used by a certain population P. Let this be a perfect case of normal language use. Imagine what would go on; and review the definition of a convention to verify that there does prevail in P a convention of truthfulness and trust in £.

(1) There prevails in P at least a regularity of truthfulness and trust in £. The members of P frequently speak (or write) sentences of £ to one another. When they do, ordinarily the speaker (or writer) utters one of the sentences he believes to be true in £; and the hearer (or reader) responds by coming to share that belief of the speaker's (unless he already had it), and adjusting his other beliefs accordingly.

(2) The members of P believe that this regularity of truthfulness and trust in £ prevails among them. Each believes this because of his experience of others' past truthfulness and trust in £.

(3) The expectation of conformity ordinarily gives everyone a good reason why he himself should conform. If he is a speaker, he expects his hearer to be trusting in £; wherefore he has reason to expect that by uttering certain sentences that are true in £ according to his beliefs — by being truthful in £ in a certain way — he can impart certain beliefs that he takes to be correct. Commonly, a speaker has some reason or other for wanting to impart some or other correct beliefs. Therefore his beliefs and desires constitute a practical reason for acting in the way he does: for uttering some sentence truthfully in £.

As for the hearer: he expects the speaker to be truthful in £, wherefore he has good reason to infer that the speaker's sentence is true in £ according to the speaker's beliefs. Commonly, a hearer also has some or other reason to believe that the speaker's beliefs are correct (by and large, and perhaps with exceptions for certain topics); so it is reasonable for him to infer that the sentence he has heard is probably true in £. Thus his beliefs about the speaker give him an epistemic reason to respond trustingly in £.

We have coordination between truthful speaker and trusting hearer. Each conforms as he does to the prevailing regularity of truthfulness and trust in £ because he expects complementary conformity on the part of the other.

But there is also a more diffuse and indirect sort of coordination. In coordinating with his present partner, a speaker or hearer also is coordinating with all those whose past truthfulness and trust in £ have contributed to his partner's present expectations. This indirect coordination

is a four-way affair: between present speakers and past speakers, present speakers and past hearers, present hearers and past speakers, and present hearers and past hearers. And whereas the direct coordination between a speaker and his hearer is a coordination of truthfulness with trust for a single sentence of £, the indirect coordination with one's partner's previous partners (and with *their* previous partners, etc.) may involve various sentences of £. It may happen that a hearer, say, has never before encountered the sentence now addressed to him; but he forms the appropriate belief on hearing it — one such that he has responded trustingly in £ — because his past experience with truthfulness in £ has involved many sentences grammatically related to this one.

(4) There is in P a general preference for general conformity to the regularity of truthfulness and trust in £. Given that most conform, the members of P want all to conform. They desire truthfulness and trust in £ from each other, as well as from themselves. This general preference is sustained by a common interest in communication. Everyone wants occasionally to impart correct beliefs and bring about appropriate actions in others by means of sounds and marks. Everyone wants to preserve his ability to do so at will. Everyone wants to be able to learn about the parts of the world that he cannot observe for himself by observing instead the sounds and marks of his fellows who have been there.

(5) The regularity of truthfulness and trust in £ has alternatives. Let £′ be any language that does not overlap £ in such a way that it is possible to be truthful and trusting simultaneously in £ and in £′, and that is rich and convenient enough to meet the needs of P for communication. Then the regularity of truthfulness and trust in £′ is an alternative to the prevailing regularity of truthfulness and trust in £. For the alternative regularity, as for the actual one, general conformity by the others would give one a reason to conform; and general conformity would be generally preferred over slightly-less-than-general conformity.

(6) Finally, all these facts are common knowledge in P. Everyone knows them, everyone knows that everyone knows them, and so on. Or at any rate none believes that another doubts them, none believes that another believes that another doubts them, and so on.

In any case in which a language £ clearly is used by a population P, then, it seems that there prevails in P a convention of truthfulness and trust in £, sustained by an interest in communication. The converse is supported by an unsuccessful search for counterexamples: I have not

been able to think of any case in which there is such a convention and yet the language £ is clearly not used in the population P. Therefore I adopt this definition, claiming that it agrees with ordinary usage in the cases in which ordinary usage is fully determinate:

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

Such conventions, I claim, provide the desired connection between languages and language-using populations.

Once we understand how languages are connected to populations, whether by conventions of truthfulness and trust for the sake of communication or in some other way, we can proceed to redefine relative to a population all those semantic concepts that we previously defined relative to a language. A string of sounds or of marks is a *sentence of* P if and only if it is a sentence of some language £ which is used in P. It has a certain *meaning in* P if and only if it has that meaning in some language £ which is used in P. It is *true in* P *at* a world w if and only if it is true at w in some language £ which is used in P. It is *true in* P if and only if it is true in some language £ which is used in P.

The account just given of conventions in general, and of conventions of language in particular, differs in one important respect from the account given in my book *Convention*.[3]

Formerly, the crucial clause in the definition of convention was stated in terms of a conditional preference for conformity: each prefers to conform if the others do, and it would be the same for the alternatives to the actual convention. (In some versions of the definition, this condition was subsumed under a broader requirement of general preference for general conformity.) The point of this was to explain why the belief that others conform would give everyone a reason for conforming likewise, and so to explain the rational self-perpetuation of conventions. But a reason involving preference in this way must be a practical reason for acting, not an epistemic reason for believing. Therefore I said that conventions were regularities in action alone. It made no sense to speak of believing something in conformity to convention. (Except in the peculiar case that others' conformity to the convention gives one a practical reason to conform by acting to somehow produce a belief in oneself; but I knew that this case was irrelevant to ordinary language use.) Thus I was cut off from what I now take to be the primary sort of conventional coordination in language use: that between the action of the truthful speaker and the responsive believing of his trusting hearer. I resorted to two different substitutes.

[3] Cambridge, Mass.: Harvard University Press, 1969. A similar account was given in the original version of this paper, written in 1968.

Sometimes it is common knowledge how the hearer will want to act if he forms various beliefs, and we can think of the speaker not only as trying to impart beliefs but also as trying thereby to get the hearer to act in a way that speaker and hearer alike deem appropriate under the circumstances that the speaker believes to obtain. Then we have speaker-hearer coordination of action. Both conform to a convention of truthfulness for the speaker plus appropriate responsive action by the hearer. The hearer's trustful believing need not be part of the content of the convention, though it must be mentioned to explain why the hearer acts in conformity. In this way we reach the account of "signaling" in *Convention*, chapter IV.

But signaling was all-too-obviously a special case. There may be no appropriate responsive action for the hearer to perform when the speaker imparts a belief to him. Or the speaker and hearer may disagree about how the hearer ought to act under the supposed circumstances. Or the speaker may not know how the hearer will decide to act; or the hearer may not know that he knows; and so on. The proper hearer's response to consider is *believing*, but that is not ordinarily an action. So in considering language use in general, in *Convention*, chapter V, I was forced to give up on speaker-hearer coordination. I took instead the diffuse coordination between the present speaker and the past speakers who trained the present hearer. Accordingly, I proposed that the convention whereby a population P used a language £ was simply a convention of truthfulness in £. Speakers conform; hearers do not, until they become speakers in their turn, if they ever do.

I think now that I went wrong when I went beyond the special case of signaling. I should have kept my original emphasis on speaker-hearer coordination, broadening the definition of convention to fit. It was Jonathan Bennett[4] who showed me how that could be done: by restat-

[4] Personal communication, 1971. Bennett himself uses the broadened concept of convention differently, wishing to exhibit conventional meaning as a special case of

ing the crucial defining clause not in terms of preference for conformity but rather in terms of reasons for conformity — practical or *epistemic* reasons. The original conditional preference requirement gives way now to clause (3): the belief that others conform gives everyone a reason to conform likewise, and it would be the same for the alternatives to the actual convention. Once this change is made, there is no longer any obstacle to including the hearer's trust as part of the content of a convention.

(The old conditional preference requirement is retained, however, in consequence of the less important clause (4). Clause (3) as applied to practical reasons, but not as applied to epistemic reasons, may be subsumed under (4).)

Bennett pointed out one advantage of the change: suppose there is only one speaker of an idiolect, but several hearers who can understand him. Shouldn't he and his hearers comprise a population that uses his idiolect? More generally, what is the difference between (a) someone who does not utter sentences of a language because he does not belong to any population that uses it, and (b) someone who does not utter sentences of the language although he does belong to such a population because at present — or always, perhaps — he has nothing to say? Both are alike, so far as action in conformity to a convention of truthfulness goes. Both are vacuously truthful. In *Convention* I made it a condition of truthfulness in £ that one sometimes does utter sentences of £, though not that one speaks up on any particular occasion. But that is unsatisfactory: what degree of truthful talkativeness does it take to keep up one's active membership in a language-using population? What if someone just never thought of anything worth saying?

(There is a less important difference between my former account and the present one. Then and now, I wanted to insist that cases of convention are cases of predominant coincidence of interest. I formerly provided for this by a defining clause that seems now unduly restrictive: in any instance of the situation to which the convention applies, everyone has approximately the same preferences regarding all possible combinations of actions. Why *all*? It may be enough that they agree in preferences to the extent specified in my present clause (4). Thus I have left out the further agreement-in-preferences clause.)

Gricean meaning. See his "The Meaning-Nominalist Strategy," *Foundations of Language*, 10(1973):141–168.

Objections and Replies

Objection: Many things which meet the definition of a language given in the thesis — many functions from strings of sounds or of marks to sets of possible worlds — are not really possible languages. They could not possibly be adopted by any human population. There may be too few sentences, or too few meanings, to make as many discriminations as language-users need to communicate. The meanings may not be anything language-users would wish to communicate about. The sentences may be very long, impossible to pronounce, or otherwise clumsy. The language may be humanly unlearnable because it has no grammar, or a grammar of the wrong kind.

Reply: Granted. The so-called languages of the thesis are merely an easily specified superset of the languages we are really interested in. A language in a narrower and more natural sense is any one of these entities that could possibly — possibly in some appropriately strict sense — be used by a human population.

Objection: The so-called languages discussed in the thesis are excessively simplified. There is no provision for indexical sentences, dependent on features of the context of their utterance: for instance, tensed sentences, sentences with personal pronouns or demonstratives, or anaphoric sentences. There is no provision for ambiguous sentences. There is no provision for non-indicative sentences: imperatives, questions, promises and threats, permissions, and so on.

Reply: Granted. I have this excuse: the phenomenon of language would be not too different if these complications did not exist, so we cannot go too far wrong by ignoring them. Nevertheless, let us sketch what could be done to provide for indexicality, ambiguity, or non-indicatives. In order not to pile complication on complication we shall take only one at a time.

We may define an *indexical language* £ as a function that assigns sets of possible worlds not to its sentences themselves, but rather to sentences paired with possible occasions of their utterance. We can say that σ is true in £ at a world w on a possible occasion o of the utterance of σ if and only if w belongs to $£(\sigma, o)$. We can say that σ is true in £ on o (without mentioning a world) if and only if the world in which o is located — our actual world if o is an actual occasion of utterance of σ, or some other world if not — belongs to $£(\sigma, o)$. We can say that a

speaker is truthful in £ if he tries not to utter any sentence σ of £ unless σ would be true in £ on the occasion of his utterance of σ. We can say that a hearer is trusting in £ if he believes an uttered sentence of £ to be true in £ on its occasion of utterance.

We may define an *ambiguous language* £ as a function that assigns to its sentences not single meanings, but finite sets of alternative meanings. (We might or might not want to stipulate that these sets are non-empty.) We can say that a sentence σ is true in £ at w under some meaning if and only if w belongs to some member of $£(\sigma)$. We can say that σ is true in £ under some meaning if and only if our actual world belongs to some member of $£(\sigma)$. We can say that someone is (minimally) truthful in £ if he tries not to utter any sentence σ of £ unless σ is true in £ under some meaning. He is trusting if he believes an uttered sentence of £ to be true in £ under some meaning.

We may define a *polymodal language* £ as a function which assigns to its sentences meanings containing two components: a set of worlds, as before; and something we can call a *mood*: indicative, imperative, etc. (It makes no difference what things these are — they might, for instance, be taken as code numbers.) We can say that a sentence σ is indicative, imperative, etc., in £ according as the mood-component of the meaning $£(\sigma)$ is indicative, imperative, etc. We can say that a sentence σ is true in £, regardless of its mood in £, if and only if our actual world belongs to the set-of-worlds-component of the meaning $£(\sigma)$. We can say that someone is truthful in £ with respect to indicatives if he tries not to utter any indicative sentence of £ which is not true in £; truthful in £ with respect to imperatives if he tries to act in such a way as to make true in £ any imperative sentence of £ that is addressed to him by someone in a relation of authority to him; and so on for other moods. He is trusting in £ with respect to indicatives if he believes uttered indicative sentences of £ to be true in £; trusting in £ with respect to imperatives if he expects his utterance of an imperative sentence of £ to result in the addressee's acting in such a way as to make that sentence true in £, provided he is in a relation of authority to the addressee; and so on. We can say simply that he is truthful and trusting in £ if he is so with respect to all moods that occur in £. It is by virtue of the various ways in which the various moods enter into the definition of truthfulness and of trust that they deserve the familiar names we have given them. (I am deliberating stretching the ordinary usage of "true," "truthfulness,"

and "trust" in extending them to non-indicatives. For instance, truthfulness with respect to imperatives is roughly what we might call *obedience* in £.)

Any natural language is simultaneously indexical, ambiguous, and polymodal; I leave the combination of complications as an exercise. Henceforth, for the most part, I shall lapse into ignoring indexicality, ambiguity, and non-indicatives.

Objection: We cannot always discover the meaning of a sentence in a population just by looking into the minds of the members of the population, no matter what we look for there. We may also need some information about the causal origin of what we find in their minds. So, in particular, we cannot always discover the meaning of a sentence in a population just by looking at the conventions prevailing therein. Consider an example: What is the meaning of the sentence "Mik Karthee was wise" in the language of our 137th-century descendants, if all we can find in any of their minds is the inadequate dictionary entry: "Mik Karthee: controversial American politician of the early atomic age"? It depends, we might think, partly on which man stands at the beginning of the long causal chain ending in that inadequate dictionary entry.

Reply: If this doctrine is correct, I can treat it as a subtle sort of indexicality. The set of worlds in which a sentence σ is true in a language £ may depend on features of possible occasions of utterance of σ. One feature of a possible occasion of utterance — admittedly a more recondite feature than the time, place, or speaker — is the causal history of a dictionary entry in a speaker's mind.

As with other kinds of indexicality, we face a problem of nomenclature. Let a $meaning_1$ be that which an indexical language £ assigns to a sentence σ on a possible occasion o of its utterance: $£(\sigma, o)$, a set of worlds on our account. Let a $meaning_2$ be that fixed function whereby the $meaning_1$ in £ of a sentence σ varies with its occasions of utterance. Which one is a meaning? That is unclear — and it is no clearer which one is a sense, intension, interpretation, truth-condition, or proposition.

The objection says that we sometimes cannot find the $meaning_1$ of σ on o in P by looking into the minds of members of P. Granted. But what prevents it is that the minds do not contain enough information about o: in particular, not enough information about its causal history. We have been given no reason to doubt that we can find the $meaning_2$

of σ in P by looking into minds; and that is all we need do to identify the indexical language used by P.

An exactly similar situation arises with more familiar kinds of indexicality. We may be unable to discover the time of an utterance of a tensed sentence by looking into minds, so we may know the meaning$_2$ of the sentence uttered in the speaker's indexical language without knowing its meaning$_1$ on the occasion in question.

Objection: It makes no sense to say that a mere string of sounds or of marks can bear a meaning or a truth-value. The proper bearers of meanings and truth-values are particular speech acts.

Reply: I do not say that a string of types of sounds or of marks, by itself, can bear a meaning or truth-value. I say it bears a meaning and truth-value relative to a language, or relative to a population. A particular speech act by itself, on the other hand, can bear a meaning and truth-value, since in most cases it uniquely determines the language that was in use on the occasion of its performance. So can a particular uttered string of vocal sounds, or a particular inscribed string of marks, since in most cases that uniquely determines the particular speech act in which it was produced, which in turn uniquely determines the language.

Objection: It is circular to give an account of meanings in terms of possible worlds. The notion of a possible world must itself be explained in semantic terms. Possible worlds are models of the analytic sentences of some language, or they are the diagrams or theories of such models.[5]

Reply: I do not agree that the notion of a possible world ought to be explained in semantic terms, or that possible worlds ought to be eliminated from our ontology and replaced by their linguistic representatives — models or whatever.

For one thing, the replacement does not work properly. Two worlds indistinguishable in the representing language will receive one and the same representative.

But more important, the replacement is gratuitous. The notion of a possible world is familiar in its own right, philosophically fruitful, and

[5] Possible worlds are taken as models in S. Kripke, "A Completeness Theorem in Modal Logic," *Journal of Symbolic Logic*, 24(1959):1–15; in Carnap's recent work on semantics and inductive logic, discussed briefly in secs. 9, 10, and 25 of "Replies and Systematic Expositions," *The Philosophy of Rudolf Carnap*, ed. P. Schilpp; and elsewhere. Worlds are taken as state-descriptions — diagrams of models — in Carnap's earlier work: for instance, sec. 18 of *Introduction to Semantics*. Worlds are taken as complete, consistent novels — theories of models — in R. Jeffrey, *The Logic of Decision*, sec. 12.8.

tolerably clear. Possible worlds are deemed mysterious and objectionable because they raise questions we may never know how to answer: are any possible worlds five-dimensional? We seem to think that we do not understand possible worlds at all unless we are capable of omniscience about them — but why should we think that? Sets also raise unanswerable questions, yet most of us do not repudiate sets.

But if you insist on repudiating possible worlds, much of my theory can be adapted to meet your needs. We must suppose that you have already defined truth and analyticity in some base language — that is the price you pay for repudiating possible worlds — and you want to define them in general, for the language of an arbitrary population P. Pick your favorite base language, with any convenient special properties you like: Latin, Esperanto, Begriffsschrift, Semantic Markerese, or what have you. Let's say you pick Latin. Then you may redefine a language as any function from certain strings of sound or of marks to sentences of Latin. A sentence σ of a language £ (in your sense) is true, analytic, etc., if and only if £(σ) is true, analytic, etc., in Latin.

You cannot believe in languages in my sense, since they involve possible worlds. But I can believe in languages in your sense. And I can map your languages onto mine by means of a fixed function from sentences of Latin to sets of worlds. This function is just the language Latin, in my sense. My language £ is the composition of two functions: your language £, and my language Latin. Thus I can accept your approach as part of mine.

Objection: Why all this needless and outmoded hypostasis of meanings? Our ordinary talk about meaning does not commit us to believing in any such entities as meanings, any more than our ordinary talk about actions for the sake of ends commits us to believing in any such entities as sakes.

Reply: Perhaps there are some who hypostatize meanings compulsively, imagining that they could not possibly make sense of our ordinary talk about meaning if they did not. Not I. I hypostatize meanings because I find it convenient to do so, and I have no good reason not to. There is no point in being a part-time nominalist. I am persuaded on independent grounds that I ought to believe in possible worlds and possible beings therein, and that I ought to believe in sets of things I believe in. Once I have these, I have all the entities I could ever want.

Objection: A language consists not only of sentences with their mean-

ings, but also of constituents of sentences — things sentences are made of — with their meanings. And if any language is to be learnable without being finite, it must somehow be determined by finitely many of its constituents and finitely many operations on constituents.

Reply: We may define a class of objects called *grammars*. A grammar Γ is a triple comprising (1) a large finite *lexicon* of *elementary constituents* paired with meanings; (2) a finite set of *combining operations* which build larger constituents by combining smaller constituents, and derive a meaning for the new constituent out of the meanings of the old ones; and (3) a *representing operation* which effectively maps certain constituents onto strings of sounds or of marks. A grammar Γ generates a function which assigns meanings to certain constituents, called *constituents in* Γ. It generates another function which assigns meanings to certain strings of sounds or of marks. Part of this latter function is what we have hitherto called a language. A grammar uniquely determines the language it generates. But a language does not uniquely determine the grammar that generates it, not even when we disregard superficial differences between grammars.

I have spoken of meanings for constituents in a grammar, but what sort of things are these? Referential semantics tried to answer that question. It was a near miss, failing because contingent facts got mixed up with the meanings. The cure, discovered by Carnap,[6] is to do referential semantics not just in our actual world but in every possible world. A meaning for a name can be a function from worlds to possible individuals; for a common noun, a function from worlds to sets; for a sentence, a function from worlds to truth-values (or more simply, the set of worlds where that function takes the value truth). Other derived categories may be defined by their characteristic modes of combination. For instance, an adjective combines with a common noun to make a compound common noun; so its meaning may be a function from common-noun meanings to common-noun meanings, such that the meaning of an adjective-plus-common-noun compound is the value of this function when given as argument the meaning of the common noun being modified. Likewise a verb phrase takes a name to make a sentence; so its meaning may be a function that takes the meaning of the name as argument to give the

[6] "Replies and Systematic Expositions," sec. 9.v. A better-known presentation of essentially the same idea is in S. Kripke, "Semantical Considerations on Modal Logic," *Acta Philosophica Fennica*, 16(1963):83–94.

meaning of the sentence as value. An adverb (of one sort) takes a verb phrase to make a verb phrase, so its meaning may be a function from verb-phrase meanings to verb-phrase meanings. And so on, as far as need be, to more and more complicated derived categories.[7]

If you repudiate possible worlds, an alternative course is open to you: let the meanings for constituents in a grammar be phrases of Latin, or whatever your favorite base language may be.

A grammar, for us, is a semantically interpreted grammar — just as a language is a semantically interpreted language. We shall not be concerned with what are called grammars or languages in a purely syntactic sense. My definition of a grammar is meant to be general enough to encompass transformational or phrase-structure grammars for natural language[8] (when provided with semantic interpretations) as well as systems of formation and valuation rules for formalized languages. Like my previous definition of a language, my definition of a grammar is too general: it gives a large superset of the interesting grammars.

A grammar, like a language, is a set-theoretical entity which can be discussed in complete abstraction from human affairs. Since a grammar generates a unique language, all the semantic concepts we earlier defined relative to a language £ — sentencehood, truth, analyticity, etc. — could just as well have been defined relative to a grammar Γ. We can also handle other semantic concepts pertaining to constituents, or to the constituent structure of sentences.

We can define the meaning in Γ, denotation in Γ, etc., of a subsentential constituent in Γ. We can define the meaning in Γ, denotation in Γ, etc., of a *phrase*: a string of sounds or of marks representing a subsentential constituent in Γ via the representing operation of Γ. We can define something we may call the *fine structure of meaning* in Γ of a sentence or phrase: the manner in which the meaning of the sentence or phrase is derived from the meanings of its constituents and the way it is built out of them. Thus we can take account of the sense in which, for instance, different analytic sentences are said to differ in meaning.

[7] See my "General Semantics," *Synthese*, 22(1970):18–67.

[8] For a description of the sort of grammars I have in mind (minus the semantic interpretation) see N. Chomsky, *Aspects of the Theory of Syntax*, and G. Harman, "Generative Grammars without Transformation Rules," *Language*, 37(1963):597–616. My "constituents" correspond to semantically interpreted deep phrase-markers, or sub-trees thereof, in a transformational grammar. My "representing operation" may work in several steps and thus subsumes both the transformational and the phonological components of a transformational grammar.

Now the objection can be restated: what ought to be called a language is what I have hitherto called a grammar, not what I have hitherto called a language. Different grammar, different language — at least if we ignore superficial differences between grammars. Verbal disagreement aside, the place I gave to my so-called languages ought to have been given instead to my so-called grammars. Why not begin by saying what it is for a grammar Γ to be used by a population P? Then we could go on to define sentencehood, truth, analyticity, etc., in P as sentencehood, truth, analyticity, etc., in whatever grammar is used by P. This approach would have the advantage that we could handle the semantics of constituents in a population in an exactly similar way. We could say that a constituent or phrase has a certain meaning, denotation, etc., in P if it has that meaning, denotation, etc., in whatever grammar is used by P. We could say that a sentence or phrase has a certain fine structure of meaning in P if it has it in whatever grammar is used by P.

Unfortunately, I know of no promising way to make objective sense of the assertion that a grammar Γ is used by a population P whereas another grammar Γ', which generates the same language as Γ, is not. I have tried to say how there are facts about P which objectively select the languages used by P. I am not sure there are facts about P which objectively select privileged grammars for those languages. It is easy enough to define truthfulness and trust in a grammar, but that will not help: a convention of truthfulness and trust in Γ will also be a convention of truthfulness and trust in Γ' whenever Γ and Γ' generate the same language.

I do not propose to discard the notion of the meaning in P of a constituent or phrase, or the fine structure of meaning in P of a sentence. To propose that would be absurd. But I hold that these notions depend on our methods of evaluating grammars, and therefore are no clearer and no more objective than our notion of a *best* grammar for a given language. For I would say that a grammar Γ is used by P if and only if Γ is a best grammar for a language £ that is used by P in virtue of a convention in P of truthfulness and trust in £; and I would define the meaning in P of a constituent or phrase, and the fine structure of meaning in P of a sentence, accordingly.

The notions of a language used by P, of a meaning of a sentence in P, and so on, are independent of our evaluation of grammars. Therefore I take these as primary. The point is not to refrain from ever saying anything that depends on the evaluation of grammars. The point is to do so only when we must, and that is why I have concentrated on languages rather than grammars.

We may meet little practical difficulty with the semantics of constituents in populations, even if its foundations are as infirm as I fear. It may often happen that all the grammars anyone might call best for a given language will agree on the meaning of a given constituent. Yet there is trouble to be found: Quine's examples of indeterminacy of reference[9] seem to be disagreements in constituent semantics between alternative good grammars for one language. We should regard with suspicion any method that purports to settle objectively whether, in some tribe, "gavagai" is true of temporally continuant rabbits or time-slices thereof. You can give their language a good grammar of either kind — and that's that.

It is useful to divide the claimed indeterminacy of constituent semantics into three separate indeterminacies. We begin with undoubted objective fact: the dependence of the subject's behavioral output on his input of sensory stimulation (both as it actually is and as it might have been) together with all the physical laws and anatomical facts that explain it. (a) This information either determines or underdetermines the subject's system of propositional attitudes: in particular, his beliefs and desires. (b) These propositional attitudes either determine or underdetermine the truth conditions of full sentences — what I have here called his language. (c) The truth conditions of full sentences either determine or undetermine the meanings of sub-sentential constituents — what I have here called his grammar.

My present discussion has been directed at the middle step, from beliefs and desires to truth conditions for full sentences. I have said that the former determine the latter — provided (what need not be the case) that the beliefs and desires of the subject and his fellows are such as to comprise a fully determinate convention of truthfulness and trust in some definite language. I have said nothing here about the determinacy of the first step; and I am inclined to share in Quine's doubts about the determinacy of the third step.

Objection: Suppose that whenever anyone is party to a convention of truthfulness and trust in any language £, his competence to be party to

[9] W. V. Quine, "Ontological Relativity," *Journal of Philosophy*, 65(1968):185–212; *Word and Object*, pp. 68–79.

that convention — to conform, to expect conformity, etc. — is due to his possession of some sort of unconscious internal representation of a grammar for £. That is a likely hypothesis, since it best explains what we know about linguistic competence. In particular, it explains why experience with some sentences leads spontaneously to expectations involving others. But on that hypothesis, we might as well bypass the conventions of language and say that £ is used by P if and only if everyone in P possesses an internal representation of a grammar for £.

Reply: In the first place, the hypothesis of internally represented grammars is not an explanation — best or otherwise — of anything. Perhaps it is *part* of some theory that best explains what we know about linguistic competence; we can't judge until we hear something about what the rest of the theory is like.

Nonetheless, I am ready enough to believe in internally represented grammars. But I am much less certain that there are internally represented grammars than I am that languages are used by populations; and I think it makes sense to say that languages might be used by populations even if there were no internally represented grammars. I can tentatively agree that £ is used by P if and only if everyone in P possesses an internal representation of a grammar for £, if that is offered as a scientific hypothesis. But I cannot accept it as any sort of analysis of "£ is used by P", since the analysandum clearly could be true although the analysans was false.

Objection: The notion of a convention of truthfulness and trust in £ is a needless complication. Why not say, straightforwardly, that £ is used by P if and only if there prevails in P a convention to bestow upon each sentence of £ the meaning that £ assigns to it? Or, indeed, that a grammar Γ of £ is used by P if and only if there prevails in P a convention to bestow upon each constituent in Γ the meaning that Γ assigns to it?

Reply: A convention, as I have defined it, is a regularity in action, or in action and belief. If that feature of the definition were given up, I do not see how to salvage any part of my theory of conventions. It is essential that a convention is a regularity such that conformity by others gives one a reason to conform; and such a reason must either be a practical reason for acting or an epistemic reason for believing. What other kind of reason is there?

Yet there is no such thing as an action of bestowing a meaning (except

for an irrelevant sort of action that is performed not by language-users but by creators of language) so we cannot suppose that language-using populations have conventions to perform such actions. Neither does bestowal of meaning consist in forming some belief. Granted, bestowal of meaning is conventional in the sense that it depends on convention: the meanings would have been different if the conventions of truthfulness and trust had been different. But bestowal of meaning is not an action done in conformity to a convention, since it is not an action, and it is not a belief-formation in conformity to a convention, since it is not a belief-formation.

Objection: The beliefs and desires that constitute a convention are inaccessible mental entities, just as much as hypothetical internal representations of grammars are. It would be best if we could say in purely behavioristic terms what it is for a language £ to be used by a population P. We might be able to do this by referring to the way in which members of P would answer counterfactual questionnaires; or by referring to the way in which they would or would not assent to sentences under deceptive sensory stimulation; or by referring to the way in which they would intuitively group sentences into similarity-classes; or in some other way.

Reply: Suppose we succeeded in giving a behavioristic operational definition of the relation "£ is used by P." This would not help us to understand what it is for £ to be used by P; for we would have to understand that already, and also know a good deal of common-sense psychology, in order to check that the operational definition was a definition of what it is supposed to be a definition of. If we did not know what it meant for £ to be used by P, we would not know what sort of behavior on the part of members of P would indicate that £ was used by P.

Objection: The conventions of language are nothing more nor less than our famously obscure old friends, the rules of language, renamed.

Reply: A convention of truthfulness and trust in £ might well be called a rule, though it lacks many features that have sometimes been thought to belong to the essence of rules. It is not promulgated by any authority. It is not enforced by means of sanctions except to the extent that, because one has some sort of reason to conform, something bad may happen if one does not. It is nowhere codified and therefore is not "laid down in the course of teaching the language" or "appealed to in

the course of criticizing a person's linguistic performance."[10] Yet it is more than a mere regularity holding "as a rule"; it is a regularity accompanied and sustained by a special kind of system of beliefs and desires.

A convention of truthfulness and trust in £ might have as consequences other regularities which were conventions of language in their own right: specializations of the convention to certain special situations. (For instance, a convention of truthfulness in £ on weekdays.) Such derivative conventions of language might also be called rules; some of them might stand a better chance of being codified than the overall convention which subsumes them.

However, there are other so-called rules of language which are not conventions of language and are not in the least like conventions of language: for instance, "rules" of syntax and semantics. They are not even regularities and cannot be formulated as imperatives. They might better be described not as rules, but as clauses in the definitions of entities which are to be mentioned in rules: clauses in the definition of a language £, of the act of being truthful in £, of the act of stating that the moon is blue, etc.

Thus the conventions of language might properly be called rules, but it is more informative and less confusing to call them conventions.

Objection: Language is not conventional. We have found that human capacities for language acquisition are highly specific and dictate the form of any language that humans can learn and use.

Reply: It may be that there is less conventionality than we used to think: fewer features of language which depend on convention, more which are determined by our innate capacities and therefore are common to all languages which are genuine alternatives to our actual language. But there are still conventions of language; and there are still convention-dependent features of language, differing from one alternative possible convention of language to another. That is established by the diversity of actual languages. There are conventions of language so long as the regularity of truthfulness in a given language has even a single alternative.

Objection: Unless a language-user is also a set-theorist, he cannot expect his fellows to conform to a regularity of truthfulness and trust in a

[10] P. Ziff, *Semantic Analysis*, pp. 34–35.

certain language £. For to conform to this regularity is to bear a relation to a certain esoteric entity: a set of ordered pairs of sequences of sound-types or of mark-types and sets of possible worlds (or something more complicated still, if £ is a natural language with indexicality, ambiguity, and non-indicatives). The common man has no concept of any such entity. Hence he can have no expectations regarding such an entity.

Reply: The common man need not have any concept of £ in order to expect his fellows to be truthful and trusting in £. He need only have suitable particular expectations about how they might act, and how they might form beliefs, in various situations. He can tell whether any actual or hypothetical particular action or belief-formation on their part is compatible with his expectations. He expects them to conform to a regularity of truthfulness and trust in £ if any particular activity or belief-formation that would fit his expectations would fall under what *we* — but not *he* — could describe as conformity to that regularity.

It may well be that his elaborate, infinite system of potential particular expectations can only be explained on the hypothesis that he has some unconscious mental entity somehow analogous to a general concept of £ — say, an internally represented grammar. But it does not matter whether this is so or not. We are concerned only to say what system of expectations a normal member of a language-using population must have. We need not engage in psychological speculation about how those expectations are generated.

Objection: If there are conventions of language, those who are party to them should know what they are. Yet no one can fully describe the conventions of language to which he is supposedly a party.

Reply: He may nevertheless know what they are. It is enough to be able to recognize conformity and non-conformity to his convention, and to be able to try to conform to it. We know ever so many things we cannot put into words.

Objection: Use of language is almost never a rational activity. We produce and respond to utterances by habit, not as the result of any sort of reasoning or deliberation.

Reply: An action may be rational, and may be explained by the agent's beliefs and desires, even though that action was done by habit, and the agent gave no thought to the beliefs or desires which were his reason for acting. A habit may be under the agent's rational control in this sense: if that habit ever ceased to serve the agent's desires accord-

ing to his beliefs, it would at once be overridden and corrected by conscious reasoning. Action done by a habit of this sort is both habitual and rational. Likewise for habits of believing. Our normal use of language is rational, since it is under rational control.

Perhaps use of language by young children is not a rational activity. Perhaps it results from habits which would not be overridden if they ceased to serve the agent's desires according to his beliefs. If that is so, I would deny that these children have yet become party to conventions of language, and I would deny that they have yet become normal members of a language-using population. Perhaps language is first acquired and afterward becomes conventional. That would not conflict with anything I have said. I am not concerned with the way in which language is acquired, only with the condition of a normal member of a language-using population when he is done acquiring language.

Objection: Language could not have originated by convention. There could not have been an agreement to begin being truthful and trusting in a certain chosen language, unless some previous language had already been available for use in making the agreement.

Reply: The first language could not have originated by an agreement, for the reason given. But that is not to say that language cannot be conventional. A convention is so-called because of the way it persists, not because of the way it originated. A convention need not originate by convention — that is, by agreement — though many conventions do originate by agreement, and others could originate by agreement even if they actually do not. In saying that language is convention-governed, I say nothing whatever about the origins of language.

Objection: A man isolated all his life from others might begin — through genius or a miracle — to use language, say to keep a diary. (This would be an accidentally private language, not the necessarily private language Wittgenstein is said to have proved to be impossible.) In this case, at least, there would be no convention involved.

Reply: Taking the definition literally, there would be no convention. But there would be something very similar. The isolated man conforms to a certain regularity at many different times. He knows at each of these times that he has conformed to that regularity in the past, and he has an interest in uniformity over time, so he continues to conform to that regularity instead of to any of various alternative regularities that would have done about as well if he had started out using them. He

knows at all times that this is so, knows that he knows at all times that this is so, and so on. We might think of the situation as one in which a convention prevails in the population of different time-slices of the same man.

Objection: It is circular to define the meaning in P of sentences in terms of the beliefs held by members of P. For presumably the members of P think in their language. For instance, they hold beliefs by accepting suitable sentences of their language. If we do not already know the meaning in P of a sentence, we do not know what belief a member of P would hold by accepting that sentence.

Reply: It may be true that men think in language, and that to hold a belief is to accept a sentence of one's language. But it does not follow that belief should be analyzed as acceptance of sentences. It should not be. Even if men do in fact think in language, they might not. It is at least possible that men — like beasts — might hold beliefs otherwise than by accepting sentences. (I shall not say here how I think belief *should* be analyzed.) No circle arises from the contingent truth that a member of P holds beliefs by accepting sentences, so long as we can specify his beliefs without mentioning the sentences he accepts. We can do this for men, as we can for beasts.

Objection: Suppose a langauge £ is used by a population of inveterate liars, who are untruthful in £ more often than not. There would not be even a regularity — still less a convention, which implies a regularity — of truthfulness and trust in £.

Reply: I deny that £ is used by the population of liars. I have undertaken to follow ordinary usage only where it is determinate; and, once it is appreciated just how extraordinary the situation would have to be, I do not believe that ordinary usage is determinate in this case. There are many similarities to clear cases in which a language is used by a population, and it is understandable that we should feel some inclination to classify this case along with them. But there are many important differences as well.

Although I deny that the population of liars *collectively* uses £, I am willing to say that each liar *individually* may use £, provided that he falsely believes that he is a member — albeit an exceptional, untruthful member — of a population wherein there prevails a convention of truthfulness and trust in £. He is in a position like that of a madman who thinks he belongs to a population which uses £, and behaves according-

ly, and so can be said to use £, although in reality all the other members of this £-using population are figments of his imagination.

Objection: Suppose the members of a population are untruthful in their language £ more often than not, not because they lie, but because they go in heavily for irony, metaphor, hyperbole, and such. It is hard to deny that the language £ is used by such a population.

Reply: I claim that these people are truthful in their language £, though they are not *literally truthful* in £. To be literally truthful in £ is to be truthful in another language related to £, a language we can call literal-£. The relation between £ and literal-£ is as follows: a good way to describe £ is to start by specifying literal-£ and then to describe £ as obtained by certain systematic departures from literal-£. This two-stage specification of £ by way of literal-£ may turn out to be much simpler than any direct specification of £.

Objection: Suppose they are often untruthful in £ because they are not communicating at all. They are joking, or telling tall tales, or telling white lies as a matter of social ritual. In these situations, there is neither truthfulness nor trust in £. Indeed, it is common knowledge that there is not.

Reply: Perhaps I can say the same sort of thing about this non-serious language use as I did about non-literal language use. That is: their seeming untruthfulness in non-serious situations is untruthfulness not in the language £ that they actually use, but only in a simplified approximation to £. We may specify £ by first specifying the approximation language, then listing the signs and features of context by which non-serious language use can be recognized, then specifying that when these signs or features are present, what would count as untruths in the approximation language do not count as such in £ itself. Perhaps they are automatically true in £, regardless of the facts; perhaps they cease to count as indicative.

Example: what would otherwise be an untruth may not be one if said by a child with crossed fingers. Unfortunately, the signs and features of context by which we recognize non-serious language use are seldom as simple, standardized, and conventional as that. While they must find a place somewhere in a full account of the phenomenon of language, it may be inexpedient to burden the specification of £ with them.

Perhaps it may be enough to note that these situations of non-serious language use must be at least somewhat exceptional if we are to have

anything like a clear case of use of £; and to recall that the definition of a convention was loose enough to tolerate some exceptions. We could take the non-serious cases simply as violations — explicable and harmless ones — of the conventions of language.

There is a third alternative, requiring a modification in my theory. We may say that a *serious communication situation* exists with respect to a sentence σ of £ whenever it is true, and common knowledge between a speaker and a hearer, that (a) the speaker does, and the hearer does not, know whether σ is true in £; (b) the hearer wants to know; (c) the speaker wants the hearer to know; and (d) neither the speaker nor the hearer has other (comparably strong) desires as to whether or not the speaker utters σ. (Note that when there is a serious communication situation with respect to σ, there is one also with respect to synonyms or contradictories in £ of σ, and probably also with respect to other logical relatives in £ of σ.) Then we may say that the convention whereby P uses £ is a convention of truthfulness and trust in £ in serious communication situations. That is: when a serious communication situation exists with respect to σ, then the speaker tries not to utter σ unless it is true in £, and the hearer responds, if σ is uttered, by coming to believe that σ is true in £. If that much is a convention in P, it does not matter what goes on in other situations: they use £.

The definition here given of a serious communication resembles that of a signaling problem in *Convention*, chapter IV, the difference being that the hearer may respond by belief-formation only, rather than by what speaker and hearer alike take to be appropriate action. If this modification were adopted, it would bring my general account of language even closer to my account in *Convention* of the special case of signaling.

Objection: Truthfulness and trust cannot be a convention. What could be the alternative to uniform truthfulness — uniform untruthfulness, perhaps? But it seems that if such untruthfulness were not intended to deceive, and did not deceive, then it too would be truthfulness.

Reply: The convention is not the regularity of truthfulness and trust *simpliciter*. It is the regularity of truthfulness and trust in some particular language £. Its alternatives are possible regularities of truthfulness and trust in other languages. A regularity of uniform untruthfulness and non-trust in a language £ can be redescribed as a regularity of truthful-

ness and trust in a different language anti-£ complementary to £. Anti-£ has exactly the same sentences as £, but with opposite truth conditions. Hence the true sentences of anti-£ are all and only the untrue sentences of £.

There is a different regularity that we may call a regularity of truthfulness and trust *simpliciter*. That is the regularity of being truthful and trusting in whichever language is used by one's fellows. This regularity neither is a convention nor depends on convention. If any language whatever is used by a population P, then a regularity (perhaps with exceptions) of truthfulness and trust *simpliciter* prevails in P.

Objection: Even truthfulness and trust in £ cannot be a convention. One conforms to a convention, on my account, because doing so answers to some sort of interest. But a decent man is truthful in £ if his fellows are, whether or not it is in his interest. For he recognizes that he is under a moral obligation to be truthful in £: an obligation to reciprocate the benefits he has derived from others' truthfulness in £, or something of that sort. Truthfulness in £ may bind the decent man against his own interest. It is more like a social contract than a convention.

Reply: The objection plays on a narrow sense of "interest" in which only selfish interests count. We commonly adopt a wider sense. We count also altruistic interests and interests springing from one's recognition of obligations. It is this wider sense that should be understood in the definition of convention. In this wider sense, it is nonsense to think of an obligation as outweighing one's interests. Rather, the obligation provides one interest which may outweigh the other interests.

A convention of truthfulness and trust in £ is sustained by a mixture of selfish interests, altruistic interests, and interests derived from obligation. Usually all are present in strength; perhaps any one would be enough to sustain the convention. But occasionally truthfulness in £ answers only to interests derived from obligation and goes against one's selfish or even altruistic interests. In such a case, only a decent man will have an interest in remaining truthful in £. But I dare say such cases are not as common as moralists might imagine. A convention of truthfulness and trust among scoundrels might well be sustained — with occasional lapses — by selfish interests alone.

A convention persists because everyone has reason to conform if others do. If the convention is a regularity in action, this is to say that it

persists because everyone prefers general conformity rather than almost-general conformity with himself as the exception. A (demythologized) social contract may also be described as a regularity sustained by a general preference for general conformity, but the second term of the preference is different. Everyone prefers general conformity over a certain state of general non-conformity called the state of nature. This general preference sets up an obligation to reciprocate the benefits derived from others' conformity, and that obligation creates an interest in conforming which sustains the social contract. The objection suggests that, among decent men, truthfulness in £ is a social contract. I agree; but there is no reason why it cannot be a social contract and a convention as well, and I think it is.

Objection: Communication cannot be explained by conventions of truthfulness alone. If I utter a sentence σ of our language £, you — expecting me to be truthful in £ — will conclude that I take σ to be true in £. If you think I am well informed, you will also conclude that probably σ is true in £. But you will draw other conclusions as well, based on your legitimate assumption that it is for some good reason that I chose to utter σ rather than remain silent, and rather than utter any of the other sentences of £ that I also take to be true in £. I can communicate all sorts of misinformation by exploiting your beliefs about my conversational purposes, without ever being untruthful in £. Communication depends on principles of helpfulness and relevance as well as truthfulness.

Reply: All this does not conflict with anything I have said. We do conform to conversational regularities of helpfulness and relevance. But these regularities are not independent conventions of language; they result from our convention of truthfulness and trust in £ together with certain general facts — not dependent on any convention — about our conversational purposes and our beliefs about one another. Since they are by-products of a convention of truthfulness and trust, it is unnecessary to mention them separately in specifying the conditions under which a language is used by a population.

Objection: Let £ be the language used in P, and let £$^-$ be some fairly rich fragment of £. That is, the sentences of £$^-$ are many but not all of the sentences of £ (in an appropriate special sense if £ is infinite); and any sentence of both has the same meaning in both. Then £$^-$ also turns out to be a language used by P; for by my definition there prevails

in P a convention of truthfulness and trust in $£^-$, sustained by an interest in communication. Not one but many — perhaps infinitely many — languages are used by P.

Reply: That is so, but it is no problem. Why not say that any rich fragment of a language used by P is itself a used language?

Indeed, we will need to say such things when P is linguistically inhomogeneous. Suppose, for instance, that P divides into two classes: the learned and the vulgar. Among the learned there prevails a convention of truthfulness and trust in a language £; among P as a whole there does not, but there does prevail a convention of truthfulness and trust in a rich fragment $£^-$ of £. We wish to say that the learned have a common language with the vulgar, but that is so only if $£^-$, as well as £, counts as a language used by the learned.

Another case: the learned use $£_1$, the vulgar use $£_2$, neither is included in the other, but there is extensive overlap. Here $£_1$ and $£_2$ are to be the most inclusive languages used by the respective classes. Again we wish to say that the learned and the vulgar have a common language: in particular, the largest fragment common to $£_1$ and $£_2$. That can be so only if this largest common fragment counts as a language used by the vulgar, by the learned, and by the whole population.

I agree that we often do not count the fragments; we can speak of *the* language of P, meaning by this not the one and only thing that is a language used by P, but rather the most inclusive language used by P. Or we could mean something else: the union of all the languages used by substantial sub-populations of P, provided that some quite large fragment of this union is used by (more or less) all of P. Note that the union as a whole need not be used at all, in my primary sense, either by P or by any sub-population of P. Thus in my example of the last paragraph, *the* language of P might be taken either as the largest common fragment of $£_1$ and $£_2$ or as the union of $£_1$ and $£_2$.

Further complications arise. Suppose that half of the population of a certain town uses English, and also uses basic Welsh; while the other half uses Welsh, and also uses basic English. The most inclusive language used by the entire population is the union of basic Welsh and basic English. The union of languages used by substantial sub-populations is the union of English and Welsh, and the proviso is satisfied that some quite large fragment of this union is used by the whole population. Yet we would be reluctant to say that either of these unions is

the language of the population of the town. We might say that Welsh and English are the two languages of the town, or that basic English and basic Welsh are. It is odd to call either of the two language-unions a language; though once they are called that, it is no further oddity to say that one or other of them is the language of the town. There are two considerations. First: English, or Welsh, or basic English, or basic Welsh, can be given a satisfactory unified grammar; whereas the language-unions cannot. Second: English, or Welsh, or basic Welsh, or basic English, is (in either of the senses I have explained) the language of a large population outside the town; whereas the language-unions are not. I am not sure which of the two considerations should be emphasized in saying when a language is the language of a population.

Objection: Let £ be the language of P; that is, the language that ought to count as the most inclusive language used by P. (Assume that P is linguistically homogeneous.) Let $£^+$ be obtained by adding garbage to £: some extra sentences, very long and difficult to pronounce, and hence never uttered in P, with arbitrarily chosen meanings in $£^+$. Then it seems that $£^+$ is a language used by P, which is absurd.

A sentence never uttered at all is *a fortiori* never uttered untruthfully. So truthfulness-as-usual in £ plus truthfulness-by-silence on the garbage sentences constitutes a kind of truthfulness in $£^+$; and the expectation thereof constitutes trust in $£^+$. Therefore we have a prevailing regularity of truthfulness and trust in $£^+$. This regularity qualifies as a convention in P sustained by an interest in communication.

Reply: Truthfulness-by-silence is truthfulness, and expectation thereof is expectation of truthfulness; but expectation of truthfulness-by-silence is not yet trust. Expectation of (successful) truthfulness — expectation that a given sentence will not be uttered falsely — is a necessary but not sufficient condition for trust. There is no regularity of trust in $£^+$, so far as the garbage sentences are concerned. Hence there is no convention of truthfulness and trust in $£^+$, and $£^+$ is not used by P.

For trust, one must be able to take an utterance of a sentence as evidence that the sentence is true. That is so only if one's degree of belief that the sentence will be uttered falsely is low, not only absolutely, but as a fraction of one's degree of belief — perhaps already very low — that the sentence will be uttered at all. Further, this must be so not merely because one believes in advance that the sentence is probably true: one's degree of belief that the sentence will be uttered falsely must be sub-

stantially lower than the product of one's degree of belief that the sentence will be uttered times one's prior degree of belief that it is false. A garbage sentence of £+ will not meet this last requirement, not even if one believes to high degrees both that it is true in £+ and that it never will be uttered.

This objection was originally made, by Stephen Schiffer, against my former view that conventions of language are conventions of truthfulness. I am inclined to think that it succeeds as a counter-example to that view. I agree that £+ is not used by *P*, in any reasonable sense, but I have not seen any way to avoid conceding that £+ is a possible language — it might *really* be used — and that there does prevail in *P* a convention of truthfulness in £+, sustained by an interest in communication. Here we have another advantage of the present account over my original one.

Objection: A sentence either is or isn't analytic in a given language, and a language either is or isn't conventionally adopted by a given population. Hence there is no way for the analytic-synthetic distinction to be unsharp. But not only can it be unsharp; it usually is, at least in cases of interest to philosophers. A sharp analytic-synthetic distinction is available only relative to particular rational reconstructions of ordinary language.

Reply: One might try to explain unsharp analyticity by a theory of degrees of convention. Conventions do admit of degree in a great many ways: by the strengths of the beliefs and desires involved, and by the fraction of exceptions to the many almost-universal quantifications in the definition of convention. But this will not help much. It is easy to imagine unsharp analyticity even in a population whose conventions of language are conventions to the highest degree in every way.

One might try to explain unsharp analyticity by recalling that we may not know whether some worlds are really possible. If a sentence is true in our language in all worlds except some worlds of doubtful possibility, then that sentence will be of doubtful analyticity. But this will not help much either. Unsharp analyticity usually seems to arise because we cannot decide whether a sentence would be true in some bizarre but clearly possible world.

A better explanation would be that our convention of language is not exactly a convention of truthfulness and trust in a single language, as I have said so far. Rather it is a convention of truthfulness and trust

in whichever we please of some cluster of similar languages: languages with more or less the same sentences, and more or less the same truth-values for the sentences in worlds close to our actual world, but with increasing divergence in truth-values as we go to increasingly remote, bizarre worlds. The convention confines us to the cluster, but leaves us with indeterminacies whenever the languages of the cluster disagree. We are free to settle these indeterminacies however we like. Thus an ordinary, open-textured, imprecise language is a sort of blur of precise languages — a region, not a point, in the space of languages. Analyticity is sharp in each language of our cluster. But when different languages of our cluster disagree on the analyticity of a sentence, then that sentence is unsharply analytic among us.

Rational reconstructions have been said to be irrelevant to philosophical problems arising in ordinary, unreconstructed language. My hypothesis of conventions of truthfulness and trust in language-clusters provides a defense against this accusation. Reconstruction is not — or not always — departure from ordinary language. Rather it is selection from ordinary language: isolation of one precise language, or of a sub-cluster, out of the language-cluster wherein we have a convention of truthfulness and trust.

Objection: The thesis and the antithesis pertain to different subjects. The thesis, in which languages are regarded as semantic systems, belongs to the philosophy of artificial languages. The antithesis, in which language is regarded as part of human natural history, belongs to the philosophy of natural language.

Reply: Not so. *Both* accounts — just like almost any account of almost anything — can most easily be applied to simple, artificial, imaginary examples. Language-games are just as artificial as formalized calculi.

According to the theory I have presented, philosophy of language is a single subject. The thesis and antithesis have been the property of rival schools; but in fact they are complementary essential ingredients in any adequate account either of languages or of language.