AN EVOLUTIONIST APPROACH TO LANGUAGE

Ruth Garrett Millikan

January, 1979
ABSTRACT

I argue that looking for functions that explain the survival value of various language forms (e.g., words, surface syntactic structures) taken with their characteristic cooperative hearer responses, while looking also for functions that explain the survival value of the mental or neural equipments that learn to produce and to react to these language forms, is a reasonable and promising approach to the study of language and the philosophy of mind. The approach promises to help to unify the philosophy of language, showing clearly how the semantic or representational side of language and the performance or "doing" side of language are integrated. The approach leads us, among other places, to a naturalistic theory of representations (e.g., sentences, beliefs, intentions) that is a distant yet genuine relative of current "causal" and "historical" theories of knowledge and of reference.
An Evolutionist Approach to Language

I shall try to convince the reader that looking for functions that explain the survival value of various language devices—sentences, words, surface syntactic forms, etc.—especially if coordinated with a search also for functions that explain the survival value of the mental or neural equipments that produce and react to these devices, is a reasonable and promising approach to the study of language. I believe that this approach, persistently pursued, could help to unify separate traditions while casting new light upon old issues within the philosophies of language and of mind. I hope to leave the reader with some understanding of why I believe this, hence with a sympathy for the approach.

In this essay I will focus primarily upon (1) those relations among language device tokens that bind these into the same language and then into univocal types and (2) beginning sketches for a naturalistic theory of "representations"—e.g., sentences, beliefs, intentions—which theory would be a distant yet genuine relative of contemporary "historical" and "causal" theories of knowledge and of reference. These two themes may strike the reader as unrelated, and the first perhaps as unimportant. A difficulty that is often unavoidable when one tries out a new approach is that traditional groupings of problems are violated. Also, which problems appear central, which peripheral, may shift. I will try to ameliorate this difficulty by moving as quickly as possible, using broad strokes and intuitive notions, over the statement of the program and the necessary topic of univocity directly to some implications that the approach suggests for a theory of representations. One of the reasons for the survival of many language devices is that these represent. Problems connected with understanding the representing function as a natural and determinate function are generally agreed upon to be important. Clearly no approach to language that failed to comment upon this function would be worth considering. But to make my arguments at all precise, a great deal of preliminary work would have to be done. In the two appendices, some of these foundations are laid with more care.

* * * * * * * * * *

1 I am indebted, in rough chronological order, to Bernard Williams, Chris Hookway, Sam Wheeler, Joel Kupperman, Jerry Shaffer and David Schwarz for constructive criticisms and advice on earlier versions of this essay.
In the early pages of *Philosophical Investigations* Wittgenstein compares words to tools. "Think of the tools in a tool box: there is a hammer, pliers, a saw, a screwdriver, a glue-pot, nails and screws. -- The functions of words are as diverse as the functions of these objects." (para. 11) Surely he would have said the same about the functions of language devices generally -- words, surface syntactic forms, tonal inflections, etc. We might try to carry Wittgenstein's analogy further.

(a) Tools "have functions" but do not always serve these functions. Although the function of the screwdriver is driving screws it sometimes fails in this task. Moreover it is not always used even with the intention of driving a screw but, say, for prying, poking holes. Language devices "have functions" but do not always serve these functions. Although the function of the imperative mood is to produce action it sometimes fails in this task. Moreover it is not always used even with the intention of producing action but, say, insincerely, sarcastically, jokingly.

(b) It is true that the physical constitution of a tool is usually directly relevant to its function whereas, within a broad range, the physical forms of language devices usually appear to be arbitrary in relation to their functions. There is no sense in which the household screwdriver "might have" served the hammer function, but "dog" might have served the "and" function and *vice versa*. But there are exceptions to this observation, most clearly in the case of tools. The key to my front door has a shape that is quite arbitrary within limits in relation to its function of opening my front door. Most any shape would have done -- provided that the lock on my door was adjusted accordingly. And it is in just this sense that most any sound could have served the "dog" function -- provided that the mechanisms within hearers that respond to tokens of this sound were adjusted accordingly.

(c) Although the functions of tools are various in the extreme, there is a uniform manner in which any tool may be described as such: (1) describe the purpose of the tool; (2) describe how the tool works, hence also the constitution of the tool, the method of operating or handling the tool and other conditions normally requisite for the tool to perform its function. A uniform manner in which any language device may be described as such...?

The Tarski-Davidson tradition of semantics talks about certain kinds of words, describing these in terms of their effects upon the truth conditions of sentences in which they occur. The Austin-Searle tradition talks about other words and devices, "performatives" and "illocutionary force indicating devices", describing these in terms of conventional rules governing their use. The Grice-Schiffer-David Lewis tradition talks about (speaking here very roughly) indicatives and imperatives describing these in terms of nested speaker intentions. At least it is clear that no accepted manner of description applicable to every language device has yet emerged. But why should there be a
problem about a uniform manner of description in the case of language devices when there is none in the case of tools? Where does the analogy with tools break down?

When we say that a tool has a certain function, its "own" or "proper" function, which can be distinguished from (a) its actual functions, what it in fact does on various occasions of use and also from (b) the functions that various users intend it to perform on various occasions, we refer to a function, roughly, that the tool type was designed by someone to serve. Natural language devices are not (at least literally) devices once "designed by someone" to serve certain functions. "The function of language device A is to F" does not bear the same analysis as does "the function of tool A is to F".

Consider another analogy. Body organs and instinctive behaviors also "have functions". As is the case with both tools and language devices, not every token of such a device succeeds in serving its "own" or "proper" function. And we can also imagine a person intentionally using such a natural device, say one of his own organs or reflexes, to serve a purpose that does not accord with its proper function. For example, people usually use their hands and arms "as Nature intended" for grasping, manipulating, pushing or pulling etc. But a person can also use these members as matter upon which to draw, as subjects for physiological experimentation, as objects of aesthetic contemplation etc. Moreover, some of these natural devices appear, as do all language devices and some tools, to have forms that are relatively arbitrary in relation to their functions. For example, instinctive mating displays, bird songs, (other) ways of marking out territory, are quite specific for the various species yet arbitrary in form within broad limits. And, as is the case with language devices but not tools, these natural devices have not literally been "designed" by someone to serve their functions. "The functions" of these natural devices are, roughly, the functions upon which their continued reproduction or survival has depended.

Do language devices "have functions" that admit of the same kind of analysis? The functions of body organs and instinctive behaviors are radically diversified yet, when known, can be described in the same sort of uniform manner that the functions of tools can be described. Could a similar manner of description be used for language devices?

By "language devices" I mean sentences, words, surface syntactic forms, tonal inflections, stress patterns, punctuations and any other significant surface elements that a natural language may contain. We begin with two speculations. First, as is the case with other natural devices that are regularly reproduced by biological systems (devices such as hearts and instinctive behaviors) we suppose that normally a natural language device has continued to be proliferated only because it serves or has served a certain, describable, stable function or set of functions. Second, as is the case for example with performances of mating displays, speaker utterances of a language device presumably
are proliferated with a stable function only in so far as stable overt or covert reactions specifically to that device on the part of cooperating partners (the female, the hearer) are also proliferated. The device type in each normal case, we speculate, should have at least one historic function, perhaps simple, perhaps complex, that accounts for the continued proliferation both of tokens of the device and of corresponding cooperative hearer reactions.

Because the language device performs this function in cooperation not with a specific hearer's response mechanism but in cooperation with a random hearer, and the hearer's cooperative response mechanism likewise performs its function in cooperation with a random speaker's utterance, it is necessary that both the language device and its cooperative hearer mechanism be "standardized". Similarly, the mating dance of the male stickleback fish must perform its proper function in cooperation with the response mechanism built into a random female stickleback and vice versa. So the male dance and the female response must be standardized throughout the species if these are to perform their proper functions, hence survive. For this reason I will call the hypothesized function or functions of a given language device that account(s) for the continued proliferation both of speaker utterances and of stable cooperative (overt or covert) hearer responses sometimes its "stabilizing (proper) function" and sometimes its "standardizing (proper) function". Stabilization of the function of a given language form and standardization within a language community of the form that serves that function must be two sides of a coin.

We speculate that it is only in so far as a language device has a stabilizing and standardizing proper function which function it actually performs in a critical proportion of cases of actual use that it can survive (1) incidents in which the stabilizing function fails to be performed (e.g., the hearer does not reply with the proper cooperative response) and (2) secondary or parasitic uses, without extinction or change of function.

Before inquiring about the "stabilizing proper functions" of various "language devices" we should have in hand (1) a clearer notion of what a "proper function" is and (2) a criterion of identity or sameness for "language devices". These two topics are related, as we will see. A stabilizing proper function is a function performed by a cooperating pair of devices (e.g., utterance and hearer response), performance of which function helps to account for the survival or proliferation of both. In appendix B I will explain "helps account for the survival of" more carefully, offering a semi-formal definition of "proper function" as it is used in this essay. Here, however, it will suffice to think of a "proper function" as like a natural purpose, assuming that the notion "natural purpose" could, in principle, be reconstructed in naturalistic or non-intentional terms. Having a proper function or natural purpose is a property attributable, in the first instance, not to individuals or tokens but to what I shall call "reproductively established families". Being a member of a reproductively established
family is being a "token" of a certain kind of "type". Language device tokens are members of types in this sense of "type". Thus the notion "reproductively established family" is a key to understanding both what a proper function is and to understanding how language device tokens are grouped into types as, e.g., "the same word" or "the same syntactic form" (in, of course, the same language). So I must explain the notion "reproductively established family".

"Reproduction" of a character occurs within a relatively stable system which produces an item on the model of some earlier item or items. That is, natural laws operating within the context of the system causally determine some character of the second item to be similar or identical to some character of its model; if the model had been different in specifiable respects this would have caused the reproduction to differ accordingly. Mass production is not as such reproduction, for later items coming off a production line are not usually modeled upon—causally determined by the character of—earlier items on the line. Usually, only when these items have been modeled upon some original not made on the line do these have, taken one by one, reproductively established characters. Items that have a reproductively established character include chromosomes, artifacts that are not of original design and, most important in the present context, learned behaviors that have resulted either from imitation or from a trial and error procedure.

A "reproductively established family" is either (1) a family of devices having reproductively established characters derived ultimately from the same character of the same model or models\textsuperscript{2} or (2) a set of devices similar to one another and derived from members of the same reproductively established family or families in the same systematic way. For example, although my heart was not directly copied from my parents' hearts, my heart was determined by genes that were copied from my parents' genes which genes determined the structure of my parents' hearts in the same way. Hence my heart and my parents' hearts may be said to be members of the same reproductively established family. The various tokens of screech owls' screeches and various tokens of walking behavior in human infants are thus members of reproductively established families.

Under condition (1) above, various tokens of the same elementary language device (e.g., word, surface syntactic form) are members of reproductively established families. Complex language devices (phrases, sentences) are not as such members of reproductively established families.

\textsuperscript{2}If we include the original model or models among family members, in the limiting case a reproductively established family consists of two members, an original and a copy.
families. When repeated they may, of course, accord with condition (2) above, but this fact seems to have no significance for the philosophy of language.

We classify elementary language devices into types in several ways but never by reference to form (shape, sound) alone. For example, neither "the word bolt" nor "the English word 'bolt'" unambiguously describes an unique language device type. Under any normal interpretation there are several English words "bolt".

Sometimes we classify language tokens into types merely by reference to reproductively established families. Call such types "genetic types". Thus word tokens may count as examples of "the same word" even though the family of which these are members has undergone slow changes in form, say, from Early to Middle to Modern English, and even though the family has become divided by reference to function and form into many distinct branches. For example, if one looks in Webster's International Dictionary second edition one can ascertain that "mean" as in "no mean city" is the same word as "mean" in "he was mean to me" but a different word than "mean" in "I didn't mean to"; that "mean" in "I didn't mean to" is the same word as "moan" in "he moaned and groaned". Because word tokens normally are copied because they serve a certain function and in order to serve the same or a related function, for the most part word tokens of the same genetic type serve at least related functions. But the parrot that mimics tokens of the English word "hello" tokens the English word "hello" and the child who uncomprehendingly copies out "il pleut" writes down the French word "il", which shows that these categories are not here defined by reference to function. Furthermore, the Martian who due to historical accident utters what sounds just like the French "il pleut", even though he may happen to mean just what the Frenchman typically means when he used this sound, does not utter the French but rather the Martian word "il".

Compare, for example, Donald Davidson: "For languages (as Quine remarks in a similar context in Word and Object) are at least as badly indi­viduated, and for much the same reasons, as propositions. Indeed, an obvious proposal linking them is this: languages are identical when identical sentences express identical propositions." ("On Saying That", Synthese 19 (1968-69) 130-146) Many other examples might be cited of the failure to note that what makes tokens of a word or sentence tokens of the same word or sentence (in the same language) is, in the first instance, history, not function.

Syntactic forms, as well as words, may look alike without being of the same genetic type. Perhaps the hypothesis that different tokens of the same "surface structure" are derived from different "deep structures" could be interpreted as an hypothesis, not about inner psychological processes, but about the evolutionary histories of these tokens. In any event such tokens were derived or copied by the speaker from different reproductively established families. It follows that they are not really tokens of the same language structure in any significant sense, but look alike—bear the same surface description—by sheer accident.
For other purposes, we divide such sprawling families into smaller families or branches paying closer attention to function and form. Thus types corresponding to separate main dictionary entries are determined according to family and form in accordance with modern spelling. They are also very roughly divided in accordance with function under the headings "noun", "verb" etc.

Finally, word tokens falling under the same main dictionary entry are again divided according to whether they have the "same sense" or "different senses" labeled "a,b,c" or "1,2,3" etc. (These "senses" are not Fregean senses; "sense" here is not the complement of "reference". For example, several such "senses" are listed for "off", for "or" and for the adverb "still".) Call such (sub-sub-) types "univocal types". Presumably classification into univocal types for other kinds of language devices, such as syntactic forms, should be possible as well. For example, the use of the indicative mood for issuing orders in the armed services would seem to be a use of this genetic type in another "sense" from the ordinary or original. But upon what kind of feature is the classification into univocal types based?

I propose that the several senses of a word or genetic language device type, its several univocal types, are not distinguished merely according to the actual functions performed, nor according to the functions intended to be performed, by the tokens to be so classified. For example, the dictionary does not attribute several senses to a word merely because it can be used metaphorically, loosely, sarcastically, playfully, deceitfully. Rather the several senses correspond to independently sufficient stabilizing functions of the word or device type, hence to independent branches of the reproductively established families of which these are members. Each of these functions is capable by itself of accounting for continued proliferation of tokens of the device. Each accounts for the survival of an independent branch of the device-type family, tokens of each branch being currently copied from earlier tokens of the same branch rather than derived again from tokens of other branches or from earlier (often obsolete) branches. Each such function is entirely capable of continuing to stabilize and standardize the language form that performs it even if all of the other branches of the family should die out. Each such standardizing function may also be capable of surviving failing uses and supporting parasitic uses such as metaphoric use, sarcastic use. Such secondary functions differ from the stabilizing functions from which they are derived in that their performance does not, indeed usually could not, taken alone, account for or encourage sustained proliferation of that device serving that function.

In appendix A I clarify the distinction between (a) stabilizing, (b) failing and (c) parasitic tokens of language devices. I argue that some kind of solid theory is needed if one wishes to draw a non-arbitrary distinction between "primary" and "secondary" functions of language devices. I suggest that the Grice-Schiffer-David Lewis
tradition that sees language use as accompanied by webs of nested intentions is an example of troubles caused, in part, by lack of such a theory. The theory of "standardizing" or "stabilizing" functions promises to be a tool by which to understand both the distinction between new senses of a term and secondary uses of a term retaining the same sense, and the distinction and relation between primary and other functions of a language device.

If the outlines we are drawing for an interpretation of language on a biological model are correct, it follows that public language is to the idiolect rather as the biological species is to the individual. The business of the biological species of keeping in business determines standards for individuals of that species which standards, though they usually correspond to averages, are not defined in terms of mere averages over the species. Consider, for example, how few sperm, or for that matter immature fishes or mice, actually manage to perform all of the functions that, none-the-less, are proper to them—that account for their survival as a species or type. Similarly, the proper functions of natural language devices, considered as elements of public languages such as French or German, are not derived by averaging over the functions of these devices within idiolects. The idiolect is not the basic unit of analysis here. Public or common language is not merely what is in common among idiolects. The business of the language species of keeping in business imposes a standard of correctness upon idiolects that is no matter of mere averages.

The standard for a language device as it appears within an idiolect, say as a disposition to utter a certain sound under certain conditions, is, minimally, that it be such that there could exist a kind of coordinate hearer reaction that would produce results that would reinforce hence stabilize both this disposition and that reaction. The idiolect does not count as an idiolect—a language form—but as a mere set of dispositions to grunt, unless (among other requirements) it could serve, does serve, or is derived from a public language that did serve or does serve\textsuperscript{5} such stabilizing functions. This is a strong sense in which purely private languages are in principle impossible. It is not that one cannot privately choose for oneself new sounds, imagining them to correlate with old functions. It is the set of functions with which these are to correlate that cannot be stipulated arbitrarily. These functions must be genuinely possible stabilizing functions—possible public functions serving both speaker and hearer—possible given how humans and the world in which they live are constituted.

Describing the stabilizing or standardizing functions of language devices not only would be a more clearly delimited project than describing the intentions of users of such devices (which often are, for a

\textsuperscript{5}—or was produced by mechanisms the proper function of which is to produce language devices that serve stabilizing functions. (See below, p. 9ff.)
particular device, diverse in the extreme) but would automatically succeed in eliminating one level of use of intentional notions from the analysis of language. Consider the analogy with which we opened this paper. Arguably, the function that defines the category "screwdriver" is not a "Function" (reserving capital F for the mathematical sense—this will prove useful later on) of the intentions of screwdriver users. It is the function that screwdrivers have been designed (by certain people) to serve. Thus it would never occur to us to attempt to describe the screwdriver's use or function by mentioning user intentions. Its function is, simply, driving screws. If a language device has something like a natural purpose, cf., a function that it paired with coordinate hearer reactions is "designed" to serve, mention of user intentions in describing its characteristic function would be similarly unnecessary or inappropriate. But in the case of language devices that represent, this takes us only one step toward a naturalist description of function. For example, duly omitting any reference to speaker intentions, I will later argue that the central stabilizing function of the indicative sentence in its root sense is the production of a true hearer belief. And I will argue that the central stabilizing function of denotative words and phrases is to produce "mental terms" or "phrases" that are then correctly "identified" within or by the hearer. But "belief" and "true", "mental term" and "correctly identify" are paradigm intentional notions. New themes must now be introduced.

First, we postulate that beliefs, intentions and "mental terms", etc., are themselves natural devices with proper functions. We postulate that they correspond to something physiological—neural structures, energy transfer patterns or whatever—these physiological devices having their own jobs to do. The performance of these jobs, when coupled with performance of other jobs by devices that cooperate with intentions and beliefs, leads to the performance of further jobs, etc., all eventually contributing to the proliferation or survival of the species that believes and intends. In this respect we take beliefs and intentions to be similar to the various other smaller and larger parts and systems that make up the human body. Call these hypothesized physiological devices corresponding to beliefs and intentions "neural sentences". Three clarifications are needed immediately.

(a) To say that a belief or intention "is" a physiological device does not here mean that being a belief or intention is the same as having a certain physiological structure. It would be absurd to suggest, for example, that having such and such a physiological structure is the same as being a belief that Columbus discovered America! Rather, to be a belief involves having certain kinds of proper functions (which functions need, of course, to be described), and it is physiological structures or activities that have proper functions. Compare: being a can opener is not having a certain physical constitution (these constitutions vary enormously) but having been designed to open cans. Yet any given can opener is thing with a certain physical constitution. Anything that is a mating display can be given a description
in terms of its physical constitution. But it is not its physical con-
stitution that makes it to be a mating display.

(b) It may help some readers to know that I do not believe that
speculating that beliefs and intentions "are" physiological devices
with proper functions has any bearing upon "reducing" the mental to
the physical. To say exactly why would require another essay. But
for one thing, something like the Kantian transcendental move may be
legitimate. Also, note that in principle no problems flavored with
Cartesian skepticism can be treated by a naturalist analysis. The
legitimacy (but not the ultimacy) of the point of view of natural
science, including its realistic ontology, is assumed here, not de-
fended.

(c) Strictly speaking it can not be specific beliefs or intentions
that have proper functions. It would be absurd to suggest, certainly
without clarification, that a novel belief or intention had a proper
function. For having a proper function has to do with evolutionary
history. Consider a chameleon that, because it is sitting on a surface
that is brown and green mottled, is itself brown and green mottled.
Strictly speaking it is not being mottled, just so, brown and green
that has a proper function here. Perhaps no chameleon has ever before
been brown and green mottled just so. If it were possible to mix a
chameleon up by clever use of ultraviolet or infrared light so that it
turned the wrong visual colors, hence was eaten by a predator, the
clear thing to say would not be that its being these visual colors had
a proper function that was not performed. Rather, the devices respon-
sible for varying its colors according to conditions had proper func-
tions that, due to odd circumstances, they were not able to perform.
Similarly, strictly speaking it is the devices responsible for forming
beliefs and intentions that are most reasonably postulated to have
proper functions. But just as we can speak generally and roughly of
"the proper functions of the color patterns of chameleons" as being
"to make them invisible to predators" we might speak generally and
roughly of "the proper function or functions of beliefs" and wonder
what this or these are.

What then might the proper functions of beliefs be? We can reason-
ably speculate that one thing they do when functioning properly is to
combine with other beliefs and intentions (neural sentences) to produce
further beliefs and intentions. We can speculate that they do this
properly only in conformity with certain rules. Presumably some, per-
haps all, of these "proper" rules--rules past conformity with which
helps to explain the survival of the belief-and-intention-combining
mechanisms--are rules that, looked upon as rules of practical and
theoretical inference, correspond to simple rules of deductive and in-
ductive logic. Roughly, when a man makes a simple mistake in reasoning
his beliefs and/or intentions are not then performing properly. Thus we
may be able to draw a distinction between the actual inference disposi-
tions that a given belief determines in a given man and those disposi-
tions that are proper to the belief. But to note that a proper function
of beliefs is to produce other beliefs in accordance with certain rules moves us only in a circle. For, it would seem, unless we know also what other proper functions beliefs have, no insight is here gained into the survival value of beliefs at all, or into the survival value of the mechanisms that produce new beliefs from old ones in accordance with definite rules. Beliefs must also have other functions, or the production of more beliefs would have no value. Likewise for intentions.

So what do beliefs do beyond participating in inference processes? The vision now begins to blur and description becomes vague. They combine with perceptions and intentions to produce actions. (Performing properly, what kind of actions?) They combine with numerous other cooperating devices to produce hopes, fears, day-dreams, night-dreams, guilt feelings, ambitions, religious attitudes, racist attitudes, depressions, anxieties, stomach ulcers, sexual arousals, releases of hormones, of adrenalin, heart attacks, violent or valiant acts. Which of these are proper functions? Which are side effects of proper functions (as when the heart says pit-a-pat in going about its business)? Which are malfunctions? Do we have to answer these questions in order to know what beliefs are?

Compare: beyond the proper function of pumping and circulating the blood there is no end to the functions to which the properly functioning heart contributes when combined with other properly functioning physiological devices. The heart helps to bring it about that the cells get oxygen, that wastes are removed, that hormones get from one place to another; hence that all other organs including the brain can perform properly; hence that beliefs are formed, inferences made, hopes produced, sentences uttered and other actions accomplished. Two points deserve attention here.

First, even if we could spell out all of the functions that a particular properly functioning heart has a disposition to perform, under these or those conditions and qua combined with properly functioning cooperative devices, this would not cast any light upon the nature of hearts qua hearts—upon the definition of the category "heart". These functions might be quite different for different individuals and certainly are for different species. In no way could the salamander's heart, qua functioning properly along with its own cooperating devices, produce hair growth, house building or utterances of English sentences. A heart is a heart because its proper function is to pump and circulate blood. Whatever further functions are thereby made possible, for this species or that, as long as there are some such functions, is quite irrelevant. Similarly, we may suppose that a description of the defining proper functions of beliefs also comes to an end—perhaps quite soon. We would not want to say, for example, that a group of people who evolved until they no longer possessed the mechanisms that produce guilt following upon certain kinds of beliefs given relevant attitudes (assuming these mechanisms to have for us a proper function) had on that account no beliefs.
Second, even if there were point, in this context, in trying to spell out all of the functions that properly functioning beliefs may contribute to in the case of current humans, how would this help us to understand beliefs qua representations? Surely it is in large part qua representations that beliefs fall into the category "belief". How would it help us to understand the primary attributes of beliefs—truth and falsity? The heart, the kidneys and the pineal gland all have (proper) dispositions to contribute to action and to speech utterances. This does not make them "representations". We must explain what it is about a belief that makes it a representation, hence different in some essential respect from various other physiological devices.

I propose that what makes a belief a representation, which among other things it needs to be to be a belief, is not just what it does. Rather it is the kind of explanation that must be given if one is fully to understand why it is able, when it is able, to perform whatever functions are proper to it. Without suggesting that the color patterns of chameleons are representations, consider this analogy. There is no way to understand why the chameleon's color pattern can serve its proper function of making it invisible to predators if one overlooks the fact that, when it does serve this function, the color pattern matches what the chameleon sits on. Analogously, there is no way to understand why a belief is able to contribute to whatever functions it may contribute to when it does function properly—in ways that have survival value—if one overlooks the fact that, exactly in these cases, it maps actual conditions in the world in accordance with rules to which cooperating devices are adjusted. Let me clarify this, building more slowly, first introducing a notion explicitly that we have already been relying upon implicitly.

We introduce the normative category "Normal", capitalized to distinguish it from "normal" in the merely statistical sense. Those conditions under which a naturally purposive device has, historically, been performing a proper function, when these conditions have contributed directly to its proper performance are of especial importance. I call these conditions "Normal" conditions for the performance of that function by that device. (Such conditions are not always statistically normal. Perhaps the device fails to perform its proper function more often than not.) More generally, we can speak of there being a Normal explanation (sometimes several alternative ones) that has accounted for the ability of a naturally purposive device to perform its proper functions when it has done so. A full description of the function of a naturally purposive device must tell not only what the device does that has accounted for its continued proliferation but also how it does it or why it is (at least sometimes) able to do it—under what conditions, in cooperation with what other cooperating devices, etc.

Typically, a description of the Normal way in which a device fulfills a certain proper function is a description of a process, often involving a number of stages. Usually, any stage in this process has the proper
function of producing any further stage. And for any stage, its achievement is itself a proper function of the device. Usually, then, there is no such thing as the proper function of a device, even when the device has an univocal function rather than several diverse functions. A description of a device with proper functions usually describes first what I call the "focal point" of any process or processes it is the proper function of the device to initiate. It locates the last point or points before the Normal process fractures into innumerable functions performed either simultaneously or alternatively and performed largely through the mediacy of cooperating devices.

Now let me try a general theory of the nature of things that represent—of representations. Stated in this general form it will be easy to show how the theory resolves certain paradoxes or problems that any naturalistic theory of representations must confront. I will introduce the theory by noting one of these paradoxes and suggesting its solution. Later I will discuss other problems.

The kind of "representations" that I have in mind to talk about is the kind where it makes sense to speak of the representation itself as being correct or incorrect as distinct from interpretations of it being correct or incorrect. Natural signs are never "representations" in this sense. I also make it a matter of definition that a "representation" is as such "complex": the value of the whole representation is a function of the value or values of a simpler element or elements of it. Thus "ouch" is not a "representation", nor is "sit" said to one's obedient dog. Perhaps I should mention also that I have in mind only actual representations, e.g., sentence tokens that are actually uttered. Merely logically possible sentences (e.g., grammatical sentences referred to by philosophers but not used) are of as little interest here as are logically possible animal behaviors to a behavioral biologist.

We introduce our general theory of representations, first, as a solution to this paradox. On the one hand representations seem to be representations in part because they correspond in accordance with specifiable rules or mapping functions to things that they represent. For example, applying a familiar way of mapping in a less familiar way, tokens of "Annette" correspond to Annette, tokens of "the father of Annette" to the father of Annette, in accordance with the rule that "the father of x" corresponds to the father of whatever corresponds to "x". Indeed, in accordance with intuitive notions, tokens of "the father of Annette loves Annette" should, similarly, correspond to Annette's father's loving Annette (in re) and so on. But on the other hand, some actual representation tokens clearly do not correspond, in accordance with the rules or kinds of mapping functions that seem relevant, to any actual things that these "represent". Names may be vacuous, descriptions unsatisfied, sentences false. It is as though representations are things that are "supposed to" correspond to or to map other real things in accordance with certain rules but that do not always do so. But how are we to understand "supposed to" here from a natural standpoint?
One job that the classical notion "intension" performed was the job of accounting for the relevant mapping rules in cases where a representation mapped nothing actual. The representation does map something—an "intension". And the "intension" maps something too—a possibility. If this possibility should happen also to be actual, the representation has not merely "intension" but "extension", or it is "true". Difficulties with this kind of solution should not need rehearsal, especially if we can substitute for it a more intelligible one. Here we need only note that this solution is no solution for a naturalist. The naturalist is interested only in laws that govern actual things, only in mapping relations that map actual relata.

Consider an analogous problem. Kidneys, no matter how variously they may be constituted in the various species, seem to be kidneys in part because they perform the function of removing wastes from the blood. On the other hand, some actual kidney "tokens" do not succeed in removing such wastes from the blood. It is as though kidneys are things that are "supposed to" remove wastes from the blood. But here we are on steadier ground. For we think we may be able to give a naturalist analysis of "supposed to" here by reference to "survival value" and evolutionary history. A somewhat better analogy, I have suggested, is that the color patterns of chameleons are "supposed to" match the color patterns of what they sit on.

Now for the trial theory of representations that has been promised. Suppose that there exists a natural device (or set of natural devices) that, when functioning in all ways properly and Normally, cooperates with a second device (or set of devices) in serving a certain function or certain kinds of functions. When these devices cooperate to perform these functions in accordance with a fully Normal explanation, the first device invariably produces and the second device responds to things that do or will map, in accordance with specifiable uniform rules, real (past, present or future) things or states of affairs or events that are in a specifiable relation to these maps. Further, it is part of the normal causal-order or natural-order explanation of how these devices, given how they are constituted or programmed, manage to perform their functions in such a manner as to encourage the survival or proliferation of both—a part of the explanation that cannot coherently be omitted—that these maps that have been produced do or will map precisely these real things so related to them in accordance with these rules. These maps, we speculate, are called "representations". They are called "representations" whether or not in a given case they are produced or reacted to properly and Normally, and whether or not they in fact map in the given case by the rules required in order for the two devices to perform properly and Normally. Those representations that do map or come to map in accordance with the rules, whether or not these have been produced or reacted to properly and Normally, are said to be "satisfied" or, in some cases, "correct" or "true". The status of the mapping rule in accordance with which a representation is "supposed to" map conditions in the real world is thus described in naturalist terms.
Consider, for example, a true sentence token. It maps, in accordance with the mapping rules Normal for it something that is in the Normal relation to it. Thus it could aid cooperating devices in the hearer that respond to it in performance of their proper functions in accordance with a fully Normal explanation. The false sentence, if it manages to aid cooperative hearer devices in performance of their proper functions, does so abnormally. It does so in accordance with an explanation, if an explanation can be given at all, that has nothing to do, historically, with the survival of these cooperating devices. More specifically (concentrating upon the hearer), first, it is an important part of the explanation for the survival of the inborn devices in a hearer that have "learned" or "been adjusted" so as to make use of sentences in a given language that these devices are capable of exactly that kind of learning. Second, it is an essential part of the Normal explanation for the fact that the current adjustment of the hearer mechanisms that produce interpretation of sentences has become stabilized, that this adjustment has in enough cases enabled proper performances on the parts of various further cooperating devices within the hearer. Suppose that the learning mechanisms have performed so as to stabilize the hearer's interpretations in accordance with a Normal explanation. Then there will be a Normal explanation for the fact that past proper performances on the part of these further cooperating devices in the hearer, given the (now) stabilized adjustments of the hearer's interpreting mechanisms, were possible. And that Normal explanation must refer to the mapping rule that has correlated past sentences reacted to "successfully" with certain configurations in the world.

Taking another example, when a simple public term is vacuous ("phlogiston") it is "a representation" (or when put in the context of a sentence becomes part of "a representation") only because it has been produced by mechanisms the proper function of which is to produce terms that are constituents of representations and, it may be important to add, because it has been produced in a way that was in large part Normal. But the process that produces a vacuous public or private term is

Normal relations between that which is mapped by a representation and the representation itself fall into two rough kinds. Where the representation contains an indexical element, the relation is or includes a causal-order relation—temporal, spatial, geometrical, causal. Where the representation contains no indexical element, the relation is "logical"—e.g., the relation of being the only real thing that the representation maps in accordance with the rules. —Or, being more realistic, being the only real thing that is both mapped and within that portion of the world frequented in thought by speaker and hearer. For we are concerned here with explanations that have actually accounted for survival of the devices that produce and respond to representations.
abNormal in one crucial respect. There is no definite thing that, because the term has often corresponded to it (in the context of complete representations and in accordance with a definite mapping function) the term has survived or its usage become or remained stabilized. Its proliferation is not explained in the same way that proliferations of public denotative terms were explained in the bulk of those cases that accounted historically for the survival of the cooperative devices that learn to produce and to react to denotative language elements. Relative to this historical explanation, its survival, so long as it does survive, is an "accident".

The newest puzzles that have been posed about representations are the various "indeterminacy" paradoxes that have been discussed at length since Quine's Word and Object. This subject can hardly be touched in a paragraph! But I can indicate something of the role that the general theory of representations I suggest—call it the "evolutionary theory of representations"—might play in relieving certain tensions. If it could be shown that what a given representation maps or is "supposed to" map in the world is, often, a determinate what, presumably the idea of determinate translation would become coherent. The problem of moving from evidence to a concrete theory of translation in a given case would only be the sort of problem (often difficult) faced by ordinary applied science in applying a general theory to specific data. Exactly what would and what would not be "preserved" by such "determinate" translation will be discussed immediately below. Here it is enough to note that the evolutionary theory of representations avoids two main sources of indeterminacy that have been shown to produce paradox. The first such source is the attempt to derive the mapping rules that relate the sentences in one language to their correct translations in another language as, merely, logical or mathematical mapping rules. One set may always be mapped upon another set in accordance with an indefinite number of mathematical mapping rules, yielding different correlations among members of these sets. The evolutionary theory of representations, however, specifies the original mapping rules—the rules that map representations onto representeds—from which original mapping rules translation rules are derived, as rules that have had, historically, something like causal efficacy. They are rules that must be referred to in giving a causal-order explanation of how interpreter devices that have reacted as they have to ancestor representations that mapped so were able to contribute to performance of certain functions. The evolutionary theory of representations is in this respect a "causal" and also an "historical" theory of representations. A second source of the indeterminacy paradoxes is the lack of a fully determinate criterion by which to distinguish "correct" from "incorrect" representations. An obvious constraint that should be imposed upon specific theories of translation is that true sentences should correspond to true sentences, false sentences to false sentences. But in the absence of any general theory of

truth (beyond, say, that most sentences that people are willing to as- sent to are "true") this constraint is not very constraining. Not only do any two sets of representations generate numerous mapping rules, any one of which might correlate members of these sets; the mem- bership of the sets to be correlated is indeterminate as well. But the evolutionary theory of representations, if it could succeed in showing that the mapping rule that is "supposed to" correlate a repre- sentation with a represented is determinate in Normal cases, would suc- ceed also in defining the notions "true" and "false" such that these apply determinately in Normal cases.

A third problem that any adequate theory of representations must deal with is understanding how the same may be represented in various dif- ferent "ways" or "manners" as distinct from being represented, merely, by materially different representations. Thus, if Jean is the father of Annette and also the father of Brigitte, "the father of Annette", "the father of Brigitte" and "Jean" represent the same (man) but in different "manners". In the Fregean tradition, these three repre- sentations are said (in accordance with a traditional translation) to have different "senses" but the same "referent". A good translation preserves not only reference but sense.

Now bees, as is well known, sometimes perform bee dances. These dances map the approximate quantity and location of honey that the dancing bee has found. There are different dialects of bee dances so that bees of one sub-species cannot necessarily interpret the dances of another species. In fact all of these dances map the place of the honey by mapping it relative to sun and hive. Suppose, however, that there were one last sub-species of bee that danced only at night, mapping the place of honey relative to the moon and the last place honey was gathered. We might then say that all of the bee sub- species but one mapped the position of honey in the same "manner". That is, all of the sub-species but one map the position of honey by reference to the same coordinate system. The last sub-species uses a different coordinate system.

If what is meant by "how a representation represents" is something like "coordinate system used", there is no difficulty in explaining how the same can be represented in different manners on the evolution- ary theory of representations. The value Jean is determined as the value of one Function in "the father of Annette", as the value of another Function in "the father of Brigitte" and as the value of an independent variable, e.g., "x" in "x loves Annette", in the repre- sentation "Jean loves Annette". So far so good. The bee dances of all sub-species but one are intertranslatable preserving sense. Dances of the last sub-species can be translated into other dances but preserving only "reference". In so far as languages do not contain corresponding variants for corresponding variables, or do not contain corresponding Functions ("father of ______"), they are not intertranslatable pre- serving sense.
But some contemporary theorists have demanded more of a good translation than that it preserve sense and reference. They have demanded, if I may follow Putnam's usage in "The Meaning of Meaning"\(^8\) that "intension" be preserved in translation. Here an "intension" is something like the program for the use of a term or sentence that is in the head of a certain speaker. These theorists have demanded that a good or ideal translation of a sentence preserve all dispositions that the speaker has to arrive at the use of that sentence \textit{via} inference from prior beliefs and ultimately \textit{via} response to perceptual data.

I have argued that what defines a belief as a representation is not to be found by looking to what it might produce—even to what it might produce when functioning \textit{properly}. For a good reason I did not attempt to show, with a parallel argument, that (part of) what defines a belief as a representation can not be found by examining what might have produced it, or what might have produced it if the belief forming mechanisms of the believer were functioning properly under conditions Normal for them. I have proposed that it is an essential part of the Normal explanation for the proper cooperative performance, hence survival, of devices or stabilized adjustments of devices that produce and react to given representations that these representations do map as they do. But, indeed, there must also be a Normal explanation, or set of alternative Normal explanations, that accounts in each case of Normal representation for the fact that the representation \textit{does} map so. In accordance with the form that this second explanation takes, Normal representations may be roughly divided into two kinds. Some come to map because they are, roughly, \textit{caused}, via the mediacy of properly adjusted representing producing devices performing under historically Normal conditions, to have the form they have by that which they represent.\(^9\) Some come to map because they \textit{cause}, via the medium of properly adjusted responding mechanisms acting under historically Normal conditions, what they represent. Normal beliefs and indicative sentences are representations of the first kind; Normal intentions and imperative sentences are representations of the second kind. That which a belief represents is a "Normal cause" of the belief. (The evolutionary theory of representations hooks up with a certain kind of "causal theory of knowledge".) Similarly, what an intention represents is a "Normal effect" of the intention.


\(^9\)—or they are caused by things that also \textit{cause} what they represent, or their form is in some other way connected by natural necessity with what they represent, e.g., \textit{via} laws of (applied) geometry.
But if the evolutionary theory of representations is roughly correct, there is no reason to suppose that an indicative sentence in a public language corresponds or ought to correspond to some particular process or set of processes that intervene between the represented as cause and the sentence as effect. First, the process of production of a particular sentence token is sometimes abNormal—because outer conditions of production were abNormal (c.f., the "Gettier examples"), or the belief producing devices were not operating properly or their current adjustments were not stable in accordance with a Normal explanation, etc. (This is always the case with false and with vacuous beliefs.) Second, there is no reason at all to suppose that there is only one kind of adjustment of a speaker's belief and sentence producing mechanisms that could (1) result from fully Normal learning processes and (2) reliably produce the same true sentences in accordance with Normal explanations. It is clear that different people may require quite different evidence before arriving quite properly and Normally, at the same belief. One piano tuner hears that the fifth he strikes is correctly tempered; another piano tuner must use instruments that the first cannot interpret. The blind and the deaf often arrive at the same true beliefs that others do but using methods that are statistically quite unusual. Just as there are lots of ways to open a can, there are often lots of alternative ways to produce a given true sentence in accordance with a Normal explanation.

There is no reason to suppose, then, that a full description of the function of any term or sentence qua in a public language should make reference to any or all specific methods whereby that term or sentence might be properly produced. Such a description of function need only note that, Normally, a term or sentence is produced in part as a result of which it represents, and via the mediacy of certain devices that have been adjusted through learning processes of a certain kind. But if there is no such thing as the dispositions in accordance with which a public language sentence is properly produced—no such thing as the intension of a public-language sentence or term—then translation from one language into another preserving intensions is indeed impossible. The idea that translation between languages might preserve more than is preserved by "homophonic translation" of idiolects within the same language is absurd.

Translation, we know, is supposed to preserve "meaning". The problem is that the notion "meaning" is a vacillating and highly ambiguous notion until placed in a specific context of inquiry. Meaning is referring or denoting.\(^{10}\) Meaning is sense. Meaning, especially for non-representational language elements, is stabilizing function. Meaning is secondary or parasitic function (see Appendix A). Meaning is speaker

\(^{10}\)See, for example, the verb "mean" in Webster's International Dictionary, second edition.
intention or purpose. Meaning is (speaker) intension. The relations among these various kinds of meaning are complex and need a full analysis. Good literal translation from one public language to another preserves reference, sense and stabilizing function. Good translation preserving the point of all secondary uses of language devices is something else. Good explications of what individual speakers mean—of intention or of intension—are two more things. (Concern with speaker intensions occurs mainly with regard to pathological cases.) What dictionaries do is still a fifth thing.

Clearly I have both said too much too fast and too little about "the evolutionary theory of representations". The point of this essay is not to propose final solutions to problems, but to bring the reader to see a sort of vision that I see—be it real or illusory. I want to open up for a serious consideration a certain approach to trying to understand language. It is not that I am sure it is the or even a correct approach. Rather, I believe that it deserves a hearing. In this context a hearing is a testing—a sustained effort to make it work—from which we might find out what it can do and what its limitations are.

An attractive feature of the approach to language that looks first for survival value, proper functions and Normal explanations in examining language devices is this. It suggests that quite diverse kinds of questions that have been asked about language devices by quite diverse traditions may be understood as not just different and each roughly valid, but as genuinely complementary, each contributing to a unified vision of language. For example, the representational value of a language device (when it has one) is something that it has only qua having, also, proper functions. The program of formal or semi-formal semantics that concentrates upon representational value, mapping rules, etc., cannot be thoroughly grounded apart from programs that examine what language devices do. In the remainder of the essay and in Appendix A, I hope to make clearer how the program we are considering might help to unify the philosophy of language. My topic for the remainder of the essay proper will be the stabilizing functions of the indicative and imperative moods and of denotative words and phrases.

The stabilizing function of a language device is postulated to be a function that helps to explain both why speakers continue to reproduce the device and why hearers continue to react to it in a standard way. That is, we expect that each univocal language device type, when functioning Normally, makes some characteristic contribution that can be understood as useful to both speaker and hearer. Its ultimate utility, however, may be "general"—as the ultimate utility of the beating of the heart is "general", contributing to many necessary further processes in the body. Better, consider the kind of "general" utility that eyes have. Their characteristic function is to enable vision. But vision aids the organism that sees in a wide variety of alternative ways, this way on this occasion, that way on another. What we are seeking is what I earlier called the "focal point" of processes that the language device
contributes to when functioning Normally. We are looking for the last point of single focus before this function fractures into diverse further functions, performed either simultaneously or alternatively and with the aid of a diversity of cooperating devices.

Consider first (the dominant sense of) imperative mood sentences. It seems clear that speakers proliferate tokens of the imperative mood mainly in so far as these tokens produce conditions that "satisfy" these sentences—in so far as hearers comply. Mature speakers may also be reinforced in their use of imperatives insofar as they perceive that their hearers attempt compliance or form intentions to comply. But on the large scale this is true only because such attempts and intentions are correlated with or understood as necessary steps toward actual compliance. If no token of the imperative mood ever effected more than an abortive attempt or intention to comply with it, it is clear that speakers would soon cease to use the imperative mood as they now do. Briefly, to produce compliance is the Normal goal or point of the speaker's utterance of an imperative sentence. The various ways that compliance contributes to satisfying further interests or goals of speakers, various further reasons that speakers may have for issuing imperatives, are, of course, scattered or diverse. As for the hearer, he responds to the sentence, first, by forming an intention in accordance with a certain rule of interpretation. This intention is then combined with beliefs, perceptions, motor activities, etc., in accordance with certain Normal processes proceeding by Normal rules. As a result of this the imperative sentence becomes "satisfied". What reinforces or perpetuates the hearer's response—the use of these rules of interpretation and these ways of combining intentions with other cooperating inner devices? The last point of single focus here—a point or goal of all hearers' Normal inner responses to imperatives—seems to be, exactly, the production of the satisfaction conditions of the sentence. Looking beyond this point, various hearers are reinforced when they respond such as to comply with imperatives by a variety of rewards and for a variety of reasons. Thus, the focused standardizing function of the imperative sentence is production of those conditions that satisfy it. That is, an imperative sentence in a given language has as its stabilizing proper function the production of conditions that are in a certain Normal relation to it and mapped by it in accordance with given rules. That this result is so ordinary, I suggest, is good confirmation of the theory that entails it. "The function" of the imperative mood is to produce that action indicated by the sentence.

Some primary ways in which complying with imperatives may produce further effects that reinforce hearer acts of compliance can be listed. These ways (roughly, as at least ostensibly purposed by speakers) correspond to the so called "illocutionary acts" that are (Normally but alternatively) performed by speakers uttering imperative mood sentences. In the case of Normally functioning orders, there are sanctions within the speaker's control to be sought or avoided; in the case of Normal requests the hearer is motivated to further the speaker's interests, either as an end in itself or as a means to further ends; when advice
is delivered in the imperative, Normally (note the capital N) the speaker advocates ends that the hearer is well to adopt in his own interests; when directions are given ("turn left here", "add the sugar slowly") the hearer is already motivated toward a goal for which the imperative supplies a means, etc.

A description of some other aspects of the Normal function of imperatives in some of these alternative modes may be found in John Searle's *Speech Acts*, 3.3ff.\(^{11}\) I suggest that the rules that Searle has laid down for performance of various "illocutionary acts" correspond, in general, to a description of the stabilizing function and Normal process and conditions, or of various alternative Normal processes and conditions, associated with various language devices used in performing these acts. Searle's mistake, I believe, is to interpret these rules as rules for performance of conventional acts. For example, suppose that the paradigm case of making a request is that in which the speaker uses a language form in a sense the stabilizing function of which is to cause a hearer to do A, uses it purposing Normal function, and hopes or assumes that the hearer will comply because he is disposed to do what the speaker wants because the speaker wants it. Then the following conditions that Searle lists for the supposedly conventional act of making a request are simply conditions for Normal production and Normal proper performance of language devices used to make requests: normal "input" and "output" conditions obtain (Searle, p. 57); the propositional content of the utterance is a future act A of H; H is able to do A; S believes that H is able to do A; it is not obvious to both S and H that H will do A in the normal course of events of his own accord; S wants H to do A (p. 66). Searle might add: S expects or hopes that H will do A; the use of the language device on S's part is an attempt to get H to do A. Because these are conditions for Normal production and Normal proper performance of the devices used to make requests, and because paradigm requesting involves purposing Normal function, these must be the usual conditions under which paradigm requests are made. Sometimes the merely ostensible making of a paradigm request is also called "making a request". In such cases not all of Searle's conditions will be met. Indeed, it is not obvious to me that any of these conditions, taken alone, is necessary to "making a request" in non-paradigm cases. But if none of these conditions were met, surely the speaker would have to drop all pretence of making a paradigm request, hence would not be "making a request" at all.

Now consider indicative sentences. It seems clear that speakers proliferate tokens of the indicative mood mainly insofar as these tokens produce beliefs, either dispositional or occurrent, in hearers. Further effects that the production of such beliefs may produce are various, and of interest to speakers for a variety of reasons.

Again, certain "illocutionary acts" performed in Normal use of the indicative provide us with a rough classification of these—reminding, informing, warning, etc. These latter functions may be interpreted as further alternative stabilizing functions of the indicative since the hearer's cooperative act in listening, interpreting, believing is reinforced insofar as he has need to be reminded, informed, warned, etc. in order to pursue his affairs. But the point of single focus for the stabilizing function of the indicative lies closer in. The hearer listens, then uses a certain rule of interpretation in forming a belief because these acts have correlated in the past with the formation of true beliefs. Sometimes hearers may be reinforced in their acts of using certain rules to form beliefs from indicative sentences merely insofar as these acts produce beliefs that accord with the speaker's or the community's beliefs. But in the final analysis this is surely so because what speakers or the community believes is highly correlated with the truth. Briefly, if no true beliefs ever resulted from hearer interpretations of indicative sentences it is clear that the indicative syntactical pattern would soon cease to be used first by hearers and, as a result, by speakers in the ways it now is. A closer look suggests that the production of a true belief in the hearer is also the dominant purpose of speaker use of the indicative mood. For nearly all of the kinds of further aims that speakers have in using indicatives can be achieved only if the hearer belief produced is true. Lying is not the rule primarily because it usually would not serve the speaker's purpose. The focused stabilizing function of the indicative mood is thus the production of a true hearer belief. Again, this result is interesting because it is so ordinary. "The function" of the indicative mood is to convey information.

Last, consider the stabilizing function of denotative devices (words, phrases) as denotative. Since denotative devices occur in both Normally functioning imperatives and Normally functioning indicatives, these must serve a function that is part of the processes both of Normal production of hearer acts from heard imperatives and Normal production of beliefs from heard indicatives. This function, I suggest, is the function of being "identified" correctly.

An act of identification takes place whenever an inference occurs of a sort that could produce a properly functioning conclusion Normally only if certain terms holding specifiable places in premisses and conclusion have the same value—denote the same. The inference from "all men are mortal" and "Socrates is a man" to "Socrates is mortal", for example, is an inference process of a kind that performs properly in accordance with a Normal explanation only insofar as whatever holds the place of "men" in the first premiss and whatever holds the place of "man" in the second have the same value. Likewise for whatever holds the place of "Socrates" in the second premiss and whatever holds the place of "Socrates" in the conclusion; for whatever holds the place of "mortal" in the first premiss and whatever holds the place of "mortal" in the conclusion. During such an inference, three acts of identification take place: an act of identifying a token of "men" with a token of
"man", an act of identifying one token of "Socrates" with another, an act of identifying one token of "mortal" with another. A more general description of the act of identifying would include among processes that proceed Normally only insofar as two representational values are the same, not only inference processes but processes that involve coupling perceptions with beliefs leading to further beliefs, processes that involve coupling perceptions with intentions leading to action, etc. The simplest kind of act of identification is an act of substitution—moving, say, from "Socrates was mortal" to "Plato's most beloved teacher was mortal". The act of arriving at a belief on the basis of perception alone is an act involving acts of identification, mental language structures being substituted for perceptual data structures. The act of expressing a belief or intention in public language involves acts of identification, overt structures being substituted for covert structures. The act of interpreting a sentence likewise involves acts of identification, covert structures being substituted for overt structures. A proper function of all denotative devices, whether mental or overt, is being identified with other representations having the same values, that is, being identified correctly.

A final remark about the program I have sketched. Such a program differs essentially from programs that proceed by attempting to give necessary and sufficient conditions for use of such phrases as "refers to x", "means that p", "knows that p", "promises to A", "tells h to A", etc. A search for such definitions would be as irrelevant to the attempt to understand language on a biological model as a search for necessary and sufficient conditions for ordinary use of the term "heart" would be for the biologist. The program seeks primarily to describe proper language functions and processes and to discover Normal explanations of how these work. It may also be useful to describe certain kinds of language malfunctions and to give certain pathologies labels. But offering definitions and counter examples, etc., would be as out of place here as it is for the physiologist. For this reason a good deal of the classical and current literature in the philosophy of language speaks only indirectly to the problems and issues discussed in this essay. That is why I have tried to motivate the reader by giving a rough and intuitive sketch of what I take to be certain potentials of the program, rather than by engaging in polemics.
Appendix A

Language device tokens may be categorized into (1) stabilizing, (2) failing and (3) parasitic. Where the speaker intends the device to fail to perform its full stabilizing function but the hearer does not, or is not intended to, we have "insincere" or "lying" uses. Where the speaker intends the stabilizing function to be performed but the hearer is purposefully responsible for the failure of the device, and where neither the speaker nor the hearer intends failure, we have other categories of failure but no handy labels. None of these cases is usually considered a "secondary use" of a language device. Any such use may be a "literal" use of a language device. Thus, tokens of imperatives that do not in fact produce action, tokens of indicatives that, unknownst to the speaker, do not express truths, tokens of devices that are not heard or not understood by the hearer, as well as lying tokens, may all be "literal" tokens. (The category "literal", though extensively relied upon in the philosophical tradition, is not a very basic or helpful category for analysis of language.) "Secondary", or what I call "parasitic" uses of language devices are, characteristically, uses that, in order to succeed, require that speaker and hearer knowingly cooperate in turning the device to another than its stabilizing function or purpose. This is possible because, when functioning properly, a language device initiates a process. During secondary use it is the intent of the speaker, in cooperation with the hearer, to reproduce only an isolated portion of that process, or perhaps only a statistically normal side effect of it. For example, the interrogative mood sentence, functioning properly in its root sense, is heard, understood, momentarily reflect upon, answered correctly, the answer understood and believed. (It is only insofar as questions are often answered correctly that they continue to be asked—that the interrogative mood survives.) Two normal side effects of the properly functioning interrogative mood sentence are that the hearer comes to believe that the speaker doesn't or didn't know, and, after the hearer replies, that the speaker comes to believe that the hearer did know. Consider what portions of this full process are invoked with secondary uses of the interrogative mood to ask rhetorical questions, examination questions, sarcastic questions.

Two language devices can have the same stabilizing function without admitting of the same secondary uses. For example, two words that denote the same, because of the contexts in which these have usually been used, may produce quite different images or stereotypes or quite different attitudes in hearers without the production of these particular side effects having anything to do with reasons for the survival of these words. To produce these effects is not then part of the stabilizing functions of these devices. On the other hand, the production of certain images, stereotypes or attitudes sometimes becomes a stabilizing function of what was originally a purely denotative idiom.

Failure to recognize the category of failing uses of language devices as separate from primary or first instance uses as well as from
full fledged secondary uses has, I believe, caused the philosophy of language much trouble. Much effort has been expended trying to understand what certain kinds of language devices do—their functions. It has, of course, been recognized that language devices have first instance or root functions and also secondary or derivative functions, and that the first problem that must be solved concerns their root functions. And it has been assumed that the univocity of an univocal language device type consisted in its various root tokens serving the same function. But no theories have been offered to explain what a root function is—to explain in what sense this function is "root" or "primary" and in what sense other functions are only "derivative" or "secondary". In the absence of any solid theory by which to focus a search for the root function of a given language device, beyond that it should be a function common at least to most tokens that are not obviously secondary, the problem of describing the root function of a language device does not have a determinate solution. Lots of things are common to the function of most not-obviously-secondary tokens of any language device. Different things in common determine different groups of tokens. Many of these groups are larger than the group of tokens that function Normally, including various kinds of failing tokens of these devices. Let me defend this statement by illustration. I will use the imperative mood as my example.

I take it that when thinkers such as Grice, Schiffer and David Lewis offer analyses of what uttering an imperative sentence and "meaning it" is, or of what "telling to" is, they are interested in more than the ordinary use of arbitrarily chosen English phrases. I take it that what they are after is a description of the same thing that is the first instance or root function of the imperative mood. In any event it is so that if we set out to look for the root function of the imperative mood with no more in the way of guidelines than the idea that this function is common to most tokens of the imperative mood that are not obviously secondary, we get the same kind of results that these philosophers have gotten. Hence I intend the following to be also a comment upon that tradition, though an indirect one.

Imperative sentences, viewed naively, serve to initiate action on the part of the hearer. But of course they do not always succeed in this task. In order to include more cases under our description of the root function of the imperative mood—cases that are not obviously "secondary uses"—we naturally adopt first the view that root tokens of the imperative mood have in common only that they express a speaker's intention that a hearer do something. Thus we can include cases in which the hearer is either incompetent or recalcitrant.

But competent users of the imperative mood do not always intend that their hearers should do as directed. There are, for example, insincere uses of the imperative mood. Shall we throw these uses into the category of secondary uses? (Cf., Schiffer's claim that in such cases there
is no "telling to" but only a pretend telling to.\textsuperscript{12}) Shall we straddle the fence by labeling the offending uses "infelicitous" (following the example of Austin and Searle)? Or shall we say instead that what defines the univocity of the imperative mood is merely that in all root cases the speaker intends that the hearer believe that the speaker intends the hearer to do something, thus moving up to the first level of embedded intentions?

But then there is this case: The speaker is angry and wants his nastiness to show. He issues an outrageous or impossible imperative. He knows quite well that the hearer will not believe that he intends its execution. But he intends that the hearer will believe that he intends him to believe that he intends him to execute it (because, say, the hearer is to think, "he wants me to try and then fail"). Is this a root use of the imperative? What evidence will decide this sort of question? If no point of further theory turns on the decision, are we not free to lay down a stipulative definition of "root" use here which would either cleanly cut this case out or cleanly move up to the next higher level of embedded intentions?

But the higher the level of embedded intentions we move to in attempting to describe the function common to all root tokens of a language device the more risk we take of failing to include mundane cases of its use and of having to relegate these to the category of secondary usage. Does the eighteen month old child who uses three word imperatives successfully possess the concepts "belief" and "intention" or is he incapable of intending that his hearer believe that he intends something? Then he does not use the imperative in its root sense? Suppose that I issue an imperative to him in the full and unquestioned belief (whether correct or incorrect) that concepts of mental states are far beyond him. I do not then intend that he believe that I intend anything; I merely intend that he act so. Is my use of the imperative here a secondary use? Does my use of the imperative with the child slowly begin to become root as I begin to recognize the first evidence that he is beginning to grasp the idea of mental states?

Again, does it matter what we say here, or are we free to stipulate a certain level of embedded intentions as defining "root" use? Might it not be simpler to avoid all reference to intentional entities in our description of the univocity of the imperative mood by simply relegateing all cases in which the imperative is not carried out to the category of secondary tokens? Why, for example, couldn't Peacocke\textsuperscript{13} merely


say that the "actual language relation" for imperatives is determined when utterance of a sentence of a certain kind harboring a proposition P is prima facie evidence that the hearer will do P?

There is a separable argument that has been used to support the view that a certain amount of nested intentional structure must underlie all normal adult uses of language, and thus should be taken as essential to a description of human language as distinguished, say, from the communication schemes of animals. Taking again the case of an imperative mood device "Do A", the argument might be isolated and clarified as follows: If the hearer H of "Do A" were to believe that the speaker S did not intend that H do A (normally) he would not do A; therefore when H complies with "Do A" he must believe that S intends that H do A. Therefore further, if S expected or intended that H would believe that S did not intend H to do A, S could not rationally intend or expect H to do A. So S in intending that H do A must also intend that H believe that S intends that H do A. Moreover, if the rational H were to believe that S intended that H should believe that S did not intend H to do A (if H takes S to be rational) H would take it that S did not intend H to do A hence would not do A. Hence the rational speaker... and so on ad infinitum.

But it never follows from the fact that having the belief that P would interfere with or be incompatible with doing a thing, that in doing that thing one must have even the dispositional belief that not-P. Analogous remarks apply to intentions, expectations, etc. If I believed that Jack the Ripper was under my bed I would not crawl into bed and instantly fall asleep. In crawling so into bed I clearly do not believe "Jack is under my bed". But it does not follow that I believe "Jack is not under my bed". Indeed, I may never have heard of Jack. Likewise, from the fact that if I had reason to believe that a speaker did not intend that I comply with an imperative then likely I would not comply, it does not follow, that in normal cases of compliance with imperatives I believe that the speaker intends compliance. (Of course were I to be asked "Did the speaker intend you to comply?" I would likely reply "Yes". But then if I were asked "What is 296 plus 379?" I might reply "675", which does nothing to show that I believed that 296+379=675 before I was asked.) If the sergeant says to the infantry private, "This time when I say 'halt!' I will intend that you not halt", if the private concentrates very hard he may succeed in not halting when he hears "halt!". It does not follow that under normal circumstances the private does more than simply react when he hears "halt!". Similarly, if the speaker of a complex imperative were to expect that his hearer would form no intention to do A when he said "Do A", presumably S could not (normally) purpose when he said "Do A" that H do A. For when A is complex, forming an intention to do A is a normal prerequisite in hearers to their doing A. But it does not follow that in issuing an imperative "Do A", S normally intends or expects that H will form an intention to do A as well as intending that H do A.

I have said, "these things do not follow", not "these things are not
true". My point is that it is a contingent matter how much nested intentional structure typically goes along with use of a certain language device, not a matter admitting of a priori proof. Furthermore, there would seem to be no reason to assume in advance that the depth of layered intentions accompanying language use is at all constant from case to case. Hence there is a very good reason for excluding reference to level of nested intentional structure from any more than stipulative definition of language or of its forms. Even the language of the infant lacking concepts of the mental may be on a strict continuum with language used by highly reflective adults and cannot be dismissed in advance as obviously a secondary use of language.

The attempt to locate the univocity of an univocal language device type in something common to the actual functions of some core group of its tokens, when unaccompanied by any prior theory concerning the sense in which this group is core, is an empty enterprise. Our first job must be to produce a theory that tells us what makes any usage primary, how various other usages depend upon primary usage, and thus how to distinguish between these.

This job is what the theory of "standardizing and stabilizing proper functions" is meant to do.
Appendix B

Not every reproductively established family or branch of such a family has a proper function. But a proper function of a device token is an attribute that it has qua a member of a reproductively established family. A device may be a member of more than one reproductively established family and may possess proper functions relative to each of these, but as possessing different reproductively established characters. Thus, a sentence token derives from several reproductively established families, the words and the grammatical form of which it is composed having different concrete histories and (complementary) proper functions.

One preliminary is needed before giving the definition of "proper function". I shall say that "a character C correlated with a function F over a set S" when a pattern of occurrences of C and F over S was of a kind that would usually indicate that the occurrence of C was a cause or condition of the performance of F (or an effect of such a cause or condition). Roughly, an observer who examined only this pattern over this set could reasonably conclude (say, in accordance with Mill's method of agreement or difference) that C had likely been a cause or condition of the performance of F in these cases. When I speak of a "correlation" of a character with a function I will be making reference to such a pattern as displayed by some definite set, not by nature overall.

In the case of a member m of a reproductively established family R having the reproductively established character C, m has the function F as a proper function iff:

1. Certain ancestors of m in R performed F

2. Because there existed a direct causal connection between having the character C and performing the function F in the case of these ancestors of m, C correlated with F over a set S including these ancestors.

3. In part because C correlated with F over S, R was proliferated, hence m has C (or exists).

The intuitive idea here is that F is a proper function of m if m exists because its character is one that can perform F. (Notice how close this is to the idea that m exists in order to perform F.) First, interpret "can perform F" as "has performed F in the past". But how can it be because C has performed F in the past as opposed to, merely, because F was performed in the past by something or other, that m was produced? For example, each moment, say A, in a cyclical process (envision a motor idling) causes further moments, say B and C, to occur.\textsuperscript{14}

\textsuperscript{14}Such causing is not usually "reproduction" of moment A of course. See the description of "reproduction" p. 5 above.
But it is not because B and C are produced by A that A recurs. It is merely because B and C are produced at the right time by something. In general, how could it ever be because it is x that causes y that x recurs, as opposed to being, merely, because y occurs?

My suggestion is that when it is because x has caused y that a correlation has existed between x and y, and because that correlation existed between x and y that x was reproduced, then it does make sense to say that x was reproduced because x caused y. Again, consider hearts. Hearts have been proliferated in part as a result of blood being circulated in the organisms containing hearts. And hearts have been proliferated by systems that, effectively, experimented by occasionally producing new variations of parts such as hearts. Thus reproduced structures the presences of which were not correlated with performances of functions useful to the organism tended to be replaced, over time, with structures that did serve useful functions. Presumably it is because the presence of hearts with the structures these exhibit has correlated with the function of circulating blood that hearts have not been replaced by other structures. It is because it is hearts that pump blood that hearts have proliferated, hence that current heart tokens exist. A proper function of hearts, then, is to circulate blood.

Behaviors that result from training or from trial and error learning involving correlations of a reward with the behavior, when the behavior has in fact been a cause of the reward, have as proper functions to produce that reward. Such behaviors are also "purposive" in another way. For the mechanisms that are responsible for such learning in animals have themselves proper functions. Their proper functions are to experiment in accordance with methods that effectively find relevant positive and negative correlations between behaviors and rewards and to proliferate behaviors accordingly, hence to bring in the rewards. They serve these functions Normally only insofar as behaviors correlate with rewards not "by accident" but due to causal connections.

Behaviors that result from imitation of behaviors of others which latter behaviors have correlated, within the observation of the learner, with certain functions, have these functions as proper functions. Most language learning, presumably, is this kind of learning. Such behaviors, again, are "purposive" also in another way, the mechanisms that are responsible for such learning having proper functions similar to those of devices responsible for trial and error learning.15

15 Clearly the correlation pattern that is (partially) responsible for the existence of a device token with a proper function may be more or less proximate vs. remote from the time of the production of that token. For example, the historical correlation pattern responsible for the appendix in humans is quite remote from current appendix tokens. Thus a proper function of a device may be called "proximate" or "remote"; the proper function of our appendices is remote. If the function accounting for reproduction of its ancestors changed over time, a device may have
So defined "proper function" is a vague term but, I would argue, properly so. For it is not the logical possibilities but the actual phenomena that divide for the most part neatly into cases where the term "proper function" has clear application and cases where it clearly does not.

Several results of the definition of "proper function" should be noted. It is not necessary that a device either serve or be capable of serving any proper function of it. E.g., very few sperm actually serve their proper functions. Also "reproduction" was defined using the vague term "similarity", so some "reproductions" may be poor copies. Correlatively, members of a proper-function category are categorized by reference neither to actual constitution nor by reference to actual or possible function. For example, it is not difficult to construct examples of accidents of nature or acts of men that might (or do) produce heart tokens quite different in form from any existent hearts, heart tokens that do not in fact circulate blood, heart tokens that could not circulate blood, items with the exact form of certain existent hearts but that are not hearts, items that do circulate blood but that are not hearts, and items that could circulate blood but are not hearts.

Ruth Garrett Millikan
Philosophy Department
University of Connecticut - U54
Storrs, CT 06268

a stacked series of proper functions, each more proximate to its current tokens than the last.

Suppose that infant Johnny reproduces Father's "damn" the first time because father's tokens of "damn" have correlated with Mother's turning attention to the speaker. Then he reproduces it a second time because token #1 correlated (by Mill's method of difference) with laughter. Token #2 will have two proper functions each of which is more proximate to it than any proper function of Father's "damn's", which latter function determines still a third proper function of token #2. The sense (dictionary "sense") of Johnny's token #2 is not determined by either of Johnny's two intentions. For example, we may say of Johnny, "he does not know what it means; don't punish him" or "he does not know what he is saying". The theory of stabilizing proper functions is needed to provide the criterion by which we locate that proper function that corresponds to the sense of a token as opposed to the intent of its user.