

Innocence: A Second Idealization for Linguistics  
 Charles J. Fillmore  
 University of California at Berkeley

1. The nature of the fit between predictions generated by a theory and the phenomena within its domain can sometimes be assessed only when different sources of explanation can be isolated through one or more idealizations. One such idealization is the simplifying assumption, for the laws of Newtonian mechanics, that the physical bodies whose movements fall within their scope are (or can be treated as) dimensionless particles, not subject to distortion or friction. The empirical laws of elasticity and friction are themselves best formulated against this background idealization.

The most frequently discussed idealization in linguistics in recent years has been that of the ideal speaker/hearer in a homogeneous speech community (Chomsky 1965, p.3f). By means of this idealization, through which we have learned a distinction between competence and performance, we are in principle able to separate out of the heterogeneous and disorderly data of speech.

(i) the systematic knowledge native speakers have about their grammars,

(ii) variation in the details of such linguistic systems from person to person, and

(iii) the effects on speech of fluctuations in speakers' attentiveness to their own texts, memory breakdowns in the course of a text's planning, any of the various kinds of speech defects, and interruptions from the surrounding world.

I am going to suggest that there is a second idealization operating in linguistics, one which underlies most traditions of semantics, and which I think it would be well to bring out into the open for careful discussion. This second idealization involves what I shall call the innocent speaker/hearer. In the way that, under the familiar idealization, the general theory of linguistic competence can be thought of as more or less equivalent to a theory of language performance on the part of an ideal speaker/hearer, we might say of certain theories of semantics that they are theories of the language-understanding abilities of the innocent speaker/hearer.

I characterize the innocent language user as fol-

lows. It knows the morphemes of its language and their meanings, it recognizes the grammatical structures and processes in which these morphemes take part, and it knows the semantic import of each of these. As a decoder, or hearer, the innocent language user calculates the meaning of each sentence from what it knows about the sentence's parts and their organization. It makes no use of past calculations: each time a structure or sentence reappears, it is calculated anew. As an encoder, or speaker, the innocent language user decides what it wishes its interlocutors to do or feel or believe and constructs a message which expresses that decision as directly as possible. There are no layers of inference between what it says and what it means.

The innocent speaker/hearer is in principle capable of saying anything sayable, given enough time. That is, its semantic system satisfies Jerrold Katz's condition of Effability and hence qualifies as a full-fledged natural language system (Katz 1972, pp.18ff). But the discourse of innocents tends to be slow, boring, and pedantic.

One early statement of our idealization is in Bloomfield's discussion of sememes and epistememes. The smallest meaningful units of lexical form are morphemes, and their meanings, we learn, are sememes; the smallest meaningful units of grammatical form are tagmemes, and their meanings are epistememes. Formally any utterance can be described as a collection of lexical and grammatical forms; semantically any utterance can be described as an assembly of its sememes and epistememes<sup>1</sup> (Bloomfield 1933, pp. 166ff).

In more recent work the semantic capabilities of an innocent speaker/hearer have come to be spoken of in terms of compositionality, a term first used, I think, in Katz and Fodor's 1963 paper on the structure of a semantic theory (Katz and Fodor 1963, p.171). More recently still John Searle has equated the idea of a compositional semantics with an assumption about the determination of a sentence's "literal meaning." In his formulation of that assumption "(T)he literal meaning of a sentence is entirely determined by the meanings of its component words (or morphemes) and the syntactical rules according to which these elements are combined" (Searle 1978, p. 207).

The model of semantic competence which fits this idealization is one which contains a lexicon, a way of characterizing grammatical structures, and a set of semantic integration rules. The model is not embarrassed by ambiguity, synonymy, homonymy or vagueness. A necessary characteristic of the model is that the meaning of a sentence in a given context is a selection from a set of meanings which the sentence has out of context. Any semantic theory which treats the determination of sentence meaning in context by a meaning-constructing rather than a meaning-selecting process goes beyond the powers of the innocence model.

2. An innocent speaker/hearer can do all of the things I said it can do, but it has several important limitations:

(1) It does not know lexical idioms, that is, lexical forms whose meanings could not be determined by somebody who knew merely their morphological structure and the meanings of their constituent morphemes. Knowing JAIL and PRISON and all possible uses of the -ER suffix could not enable an innocent to figure out the difference in meaning between JAILER and PRISONER.<sup>2</sup>

(2) The innocent language user does not know phrasal idioms. If you were to go up to it and say, YOUR GOOSE IS COOKED!, it would feel worried if it had a pet goose, grateful if it had just brought a goose carcass home for dinner, or puzzled if it had no goose at all. But it would lack the idiomatic interpretation that the rest of us are able to give the expression.

(3) The innocent language user does not know lexical collocations that are not based on necessary meaning relations. If it knows the expression BLITHERING IDIOT at all, it has to assume that BLITHERING is a form of the verb BLITHER and that it can be used of anything that 'blithers.' Understanding only the meanings of the words, it has no reason to know that BLITHERING is limited to the context in which it modifies IDIOT.<sup>3</sup>

(4) It lacks the ability to judge the appropriateness of fixed expressions to specific types of situations. It has no situational associations with such expressions as THIS HURTS ME MORE THAN IT HURTS YOU or THIS IS WHERE I CAME IN, to say nothing of such semantically opaque

locutions as KNOCK ON WOOD or SPEAK OF THE DEVIL.

(5) It possesses no construal principles for metaphorical language use, nor, in fact, does it have any reason to believe that language can be used metaphorically. It is accordingly ignorant of the conventional images that provide grounding for metaphoric interpretation in its language. Suppose we induced it to try to interpret the metaphoric utterance, I'LL STAND BEHIND YOU. It has no basis for preferring the image of a person falling backwards, an image which would allow the utterance to be taken as comforting, over that of a person falling forwards, in which case the expression could be taken as threatening.<sup>4</sup>

(6) In general the innocent language user lacks any interpretive mechanisms for indirect communication, that is, for meaning one thing while saying another, or principles of text coherence that would allow it to 'read between the lines' in a text. If we can suppose that it enjoys being flattered, then it will indeed be flattered if we say to it, YOU HAVE A VERY LOVELY LEFT EYE.<sup>5</sup>

(7) The innocent one has no background of understanding for what might be called text structure. That is, it is unable to 'situate' pieces of text within slots defined for given kinds of texts. One of the clearest examples of the kind of 'situating' I have in mind is provided by a convention in Japanese letter-writing. Personal letters in Japan are expected to begin with a preamble which contains comments on the current season. The innocent, on reading at the beginning of a letter that its Japanese correspondent's garden floor is covered with leaves will not realize that this remark serves to satisfy that convention.<sup>6</sup>

Summarizing, the innocent speaker/hearer does not know about lexical idioms, phrasal idioms, lexical collocations, situational formulas, indirect communication, or the expected structures of texts of given types. The collections of things the innocent language user does not know gives us a catalogue of the kinds of uses of and responses to language that fall outside of the ideal of a pure compositional semantics. The innocence idealization is in fact frequently thought of as establishing the boundary between semantics proper and such neighboring concerns as pragmatics, rhetoric, logic, and language comprehension. I will be showing

below that in this purest form the idealization has proved to be incompatible with the territorial urges of some semantic theorists.

3. But first I need to introduce some new distinctions. The semantics of innocence is a compositional semantics, but I need to say more now about just what that means. I would like to begin by reviewing a distinction that is sometimes ignored, that between compositionality proper and motivation. An expression can be spoken of as motivated if the speakers of the language see it as having the form it has by virtue of some (possibly vaguely perceived) word-forming or phrase-forming principles. POET, POEM and POETRY appear to be constructed out of 'partly identical material in a way that reflects their semantic commonality, and such words as JAILER and PRISONER have components that speakers see as related to their meanings. On the other hand, when we say that an expression's interpretation is compositional, we mean that the expression is more than merely 'motivated'; but what is that additional element?

To be clear about them we need to make another distinction. Here, as in many areas of linguistics, I think it is important to distinguish the decoding, or hearer's point of view from the encoding, or speaker's point of view. Applying these two perspectives in the case of compositionality, we can talk about semantic transparency in the decoding case, and semantic productivity in the encoding case. An expression is semantically transparent if we can rely on compositional semantics to figure out what it means once we encounter it. A set of syntactic-semantic rules is semantically productive if by relying on them we can succeed in producing fully natural ways of saying what we mean. The distinction I am making between the two 'directions' of compositionality bears on the distinction between two senses of the English adjective IDIOMATIC. In the one case, were I to say that your speech is 'idiomatic' I would mean that it contains expressions which the innocent language user could not interpret. In the other case, in describing your speech as idiomatic, what I would mean is that it is what an accomplished native speaker would naturally say, and that means that it is not likely to be what an innocent would have chosen to say. The distinction I have just

drawn is essentially the distinction between what Adam Makkai calls idioms of decoding and idioms of encoding (Makkai 1972).

4. The problems linguists have in dealing with the innocence idealization have been in connection with fixed expressions, collocations, idioms, indirect communication, and the differences I have just been discussing regarding motivation and compositionality. Many theoretical moves that semanticists have made seem to be directed toward increasing the domain of semantics while preserving innocence. The goal is to reformulate semantic observations in such a way that the innocence idealization fits cases it didn't fit before the reformulation, thus reducing the need to look for new sources of explanation. Compositional semantics, after all, is dependable and formally easy to cope with: the more that can be brought into its scope the better off we are. Or so it is sometimes thought.

(1) One of the innocence-preserving moves that I have in mind involves the context restriction of the senses of a polysemous word. What may have looked like a lexical or phrasal idiom will turn out to fit a purely compositional semantics if we allow ourselves to say that some of the morphemes have senses that just happen to be limited to this specific context. Zellig Harris has a beautiful formulation of this principle, using BLUEBERRY as his key example. He states, "the meaning of an element in each linguistic environment is the difference between the meaning of its linguistic environment and the meaning of the whole utterance (i.e. the whole social situation). Thus the meaning of blue in blueberry might be said to be the meaning of blueberry minus the meaning of berry and of the '--- morpheme: blue here therefore does not mean simply a color, but the observable differentia of blueberries as against other berries." (Harris 1951, p. 347). Instead of saying that the word BLUEBERRY is a composite word which somebody in the history of the English-speaking people invented as a name for a particular genus of berries, a formulation which departs from innocence, we can now be pleased to realize that the word contains exactly the right morphemes and that these together, by completely regular rules, designate exactly this genus

of berries.

(2) The Hungarian semanticist László Antal preserves innocence in a far different way: he does so by insisting on a sharp distinction between meaning and content. To Antal, every morpheme has a unique meaning, and every expression composed of morphemes has a meaning exactly represented by that assembly of morphemes. Content, by contrast, appears only with sentences and texts, not with words, and can be described only by using knowledge of facts that are clearly outside of linguistics. Meaning and content, Antal says, "differ from each other in that the former is broken down into smaller parts, while content manifests itself as an undivided whole. This, he goes on, "is because the meaning of the sentence is made up of the meanings of the individual morphemes that occur in it." Antal would say that the meaning differences separating PETER LOST HIS WAY, PETER LOST HIS MIND, PETER LOST HIS JOB and PETER LOST HIS PATIENCE are to be found precisely and unambiguously in the meaning differences separating WAY, MIND, PATIENCE and JOB.<sup>7</sup> Their differences in content are not so orderly; but that, according to Antal, is not the semanticist's concern (Antal 1964, p.23).

(3) A third move for preserving innocence is that taken by Charles Hockett; it is a decision by which expressions which might appear to some people to be morphologically complex and semantically irregular turn out to be primary linguistic units and hence to offer no challenge to compositional semantics. One way of accomplishing this kind of redefinitional solution would be to extend the range of the term morpheme to include lexical and phrasal idioms. At the lexical level this would be to say that such words as REFER, PREFER, RECEDE and PRECEDE are synchronically four separate undivided morphemes whose internal structures have only etymological relevance. Hockett's choice, by contrast, was to generalize the term idiom to make it include morphemes. If an idiom is a linguistic form whose meaning is not built up out of the meanings of its constituent parts, then morphemes are idioms. Having made this terminological choice, Hockett can then claim that "any utterance consists wholly of an integral number of idioms. Any composite form which is not itself idiomatic consists of smaller forms which are." (Hockett 1958, p.173). We are left, then, with a uni-

form class of primary meaning-bearing elements, and no troublesome distinction between morphemes and idioms.

This decision appears to leave us with the problem of not being able to recognize that certain expressions are simultaneously fixed expressions and semantically motivated. One possible solution--I don't know whether this would be Hockett's solution--is to regard what I see as a descriptive problem as simply involving the distinction between a pure synchronic description of a language on the one hand and on the other hand whatever knowledge or beliefs speakers may have of the motivational basis of given linguistic forms at the time they were introduced into the language.

(4) A fourth common move to preserve innocence is one which claims a sharp distinction between knowledge about shared meanings and knowledge about the world. By distinguishing, as it is sometimes put, a dictionary and an encyclopedia, we can allow ourselves to say that the innocent speaker/hearer can know everything about the meaning of a sentence independently of knowing anything at all about what the world is like.

The relevance of such a decision to the question of innocence is that it allows a distinction between two kinds of judgments about the acceptability of sentences, the one having to do with true semantic compatibility and the other with truth or plausibility. The semantic integration principles operate by accepting well-formed semantic complexes and rejecting ill-formed ones. This task is a more cleanly determined one if the meaning vs. world distinction is maintained. The analyst is left facing a number of decisions, however, that are extremely hard to make: to borrow a favorite example of John Searle's, we might wonder what could possibly be the difference between a description of an oscilloscope and a statement of the meaning of the noun OSCILLOSCOPE.<sup>8</sup>

(5) A fifth move to preserve innocence is that of minimizing the appearance of polysemy in semantic description and formulating invariant meanings for all uses of a morpheme or word. Rather than describing the phrase CUT THE CARDS, in the sense of dividing and restacking a deck of playing cards, as an idiom, we could simply try to formulate the meaning of the verb CUT in such a way that it included all of the uncontested uses of the



verb plus the use we see in this expression.

A commitment to the formulation of invariant or 'core' meanings, a position associated in general with Dwight Bolinger (Bolinger 1977) and with respect to grammatical morphemes in the work of Roman Jakobson (Jakobson 1936) and William Diver (Diver 1964), puts its holders into an essentially unassailable position. My argument against it is that when you have captured the core meanings of everything, you have no basis for knowing which combinations of words have which meanings. The core-meaning linguist insists that morphemes have just the meanings that they have, and that people who see a problem in constructing composite meanings out of component meanings are confused about the difference between meaning and comprehension.

(6) One final innocence-preserving strategy is the one which posits a finite number of possible relations that can link together the elements that make up a compound word.<sup>9</sup> Given this decision, we can say that the noun compounds HORSE SHOES and ALLIGATOR SHOES, to use examples from Katz and Fodor (1963), are each ambiguous in many ways, and in the same ways. Each can designate shoes that are worn by the animal named, that are made out of that animal's skin, that are made in the shape of the animal, and so on. The fact that each of these has been lexicalized in English, conventionalized as a composite name with a specific assigned sense, can be taken as proper to the study of language use rather than semantics; and the fact that in detail the kinds of relationships we sense linking the parts of compounds do not appear to be neatly classifiable is merely to be taken as evidence for the abstractness of the underlying relationships (cf. Bolinger 1965, p.568).

5. The innocence idealization has served linguistics well, at least as a heuristic for making linguists aware of the various modes of signifying; and in one form or another it is a necessary core of any theory capable of coping with the reality that speakers and hearers do indeed create and comprehend novel sentences. But I feel that the desire to generalize it has backed semanticists into analytical corners that they would have done

well to stay out of. In particular, I believe that the facts which lie outside of anything the innocence model is capable of handling are so pervasive and powerful that nothing really important is gained by distorting the idealization in the way the innocence-preservers have chosen to do. I am not arguing that the idealization be abandoned; only that it should be kept pure.

There are three characteristics of the semantic systems of real speakers and hearers that seem absolutely critical in this connection.

The first of these is what might be referred to as the 'layering' of conventionality in language. With Jerry Morgan we can say that while we need to recognize from the start the conventional or arbitrary nature of the relations between elementary signs and their meanings, we must also recognize conventional pairings of contexts with meanings-to-convey-in-those-contexts, as well as conventional pairings between contexts and particular expressions by which conventionalized meanings get conveyed in those contexts (Morgan 1978).

A second characteristic is the inescapable participation of context and background in constructing the meanings of utterances in actual use. Pamela Downing (Downing 1977), Herbert and Eve Clark (Clark and Clark 1978) and Geoffrey Nunberg (Nunberg 1977) have all given us an awareness of classes of expressions-in-use which have the following properties: they are not semantically transparent, they do not have conventionally assigned meanings, and they cannot be seen as instances of metaphoring acts of the usual kind; they are expressions for whose interpretation we require a detailed understanding of the participants' shared engagement in an experiential context. Once the operation of context is made clear in these obvious cases, it becomes possible for us to ask what role it plays in cases that used to appear to satisfy the idealization. I suspect that we will sometimes be surprised.

A third important characteristic of real language use is found in what I shall call structural formulas. A language's free phrases are limited only by the grammar and what people choose to say; fixed expressions have most of their lexical and grammatical, and maybe even their prosodic, properties fixed by convention. For the structural formulas that I have in mind, the grammatical form and possibly one or two lexical items

are fixed, but the class of substitutions is open but constrained by semantic and pragmatic considerations. As an example, consider expressions of the form 'Someone plays Something to Someone's Something', as in SHE PLAYED DESDEMONA TO MY OTHELLO, which I could use in reporting shared theatrical experiences she and I have had. As a second example, consider 'X in and X out' where X can be a word designating a cyclic calendric term, as in DAY IN AND DAY OUT, YEAR IN AND YEAR OUT, etc. (cf. Kiparsky 1976). Or consider the colloquial formula 'WATCH Something Happen', a very flexible construction beginning with 'imperative' WATCH and followed by an infinitive clause indicating a surprise which fate might have in store. (Examples: I've been insisting that you're too young to carry such a big tray of fruit, and I take it away from you. Then I say, NOW WATCH ME DROP IT. We've been planning on a picnic for many days, and have just invested a lot of money in getting the supplies for it; I say, NOW WATCH IT RAIN.)<sup>10</sup>

I expect that there are lots of structural formulas like this, each with its own private semantic interpretation rules. If the number and frequency of such constructions is very great, there might some day be semanticists who feel that the standard form of compositional semantics can be undermined altogether, by having its principles absorbed into the list of pairings of such formulas and specific semantic interpretation rules. It is conceivable that the central principle of truth-conditional semantics could be introduced, in such a system, as an interpretation rule for a structural formula called 'indicative sentence.'

The argument that this last proposal is not altogether absurd must be saved for another occasion. For now let me just hope to have convinced you of the importance of distinguishing real innocence from pretended innocence.

#### Footnotes

1. Bloomfield insists, I should point out, that any effort to characterize the sememes and episememes substantively belongs outside of linguistics proper. He seems to be saying that if we knew what these things were, we would know, as linguists, what to do with them.

2. My phrase "the -ER suffix" represents an implicit synchronic judgment about the structure of these words. The etymologically sophisticated will find the JAILER/PRISONER examples unfair. The claim, however, is simply that ordinary speakers will see these words as made up of stem plus -ER.
3. There is a problem, of course, in deciding what it is to "know the meaning of" a word that has such tight collocational requirements, especially since the expression is almost never used in what might be called its "literal" meaning. Until just a few days ago, I myself believed that BLITHERING meant 'drooling.'
4. This example is borrowed from George Lakoff.
5. It does not know Oswald Ducrot's 'loi d'exhaustivité' (Ducrot 1972, p. 170) or Paul Grice's 'Quantity maxim' (Grice 1975, pp. 45ff).
6. Nor is the innocent in a position to appreciate the following facts: (i) that urgent letters in which the seasonal remarks are left out usually begin with an apology, one version of which is the word ZENRYAKU, an abbreviation of a larger expression which means 'apology for omitting the preamble'; and (ii) that the Kenkyusha Japanese-English dictionary defines ZENRYAKU as "I hasten to inform you that . . ."
7. Antal's examples include only PATIENCE and JOB, not the other two. It is clear from the context, however, that he would accept the statement I made about the four sentences.
8. There is no doubt that for a great many cases such a distinction is necessary, but there is clearly a problem of knowing where to draw the line. For a typical statement, see Leech (1974, pp. 87ff).
9. For a treatment exemplifying this approach as applied to the special class of "complex nominals," see Levi (1978).
10. These expressions cannot be thought of merely as uses of imperative sentences with the normal embedding verb WATCH, since some of these expressions contain

clausal elements that could not normally serve as complements of WATCH. I once heard an adolescent girl, who had been fretting at some length about how unimpressive her blind date was probably going to be, say NOW WATCH HIM BE REAL HANDSOME.

References:

- Antal, Laszlo (1964), Content, Meaning and Understanding. Mouton.
- Bloomfield, Leonard (1933), Language, Holt, Rinehard and Winston.
- Bolinger, Dwight (1965), "The atomization of meaning," Language 41.555-573.
- Bolinger, Dwight (1977) Meaning and Form. Longmans.
- Chomsky, Noam (1965), Aspects of the Theory of Syntax. MIT Press.
- Clark, Herbert H. and Eve V. Clark (1978), "When nouns surface as verbs" ms.
- Diver, William (1964), "The system of agency of the Latin noun," Word 20.178-196.
- Downing, Pamela (1977), "On the creation and use of English compound nouns," Language 53.810-842.
- Ducrot, Oswald (1972), Dire et ne pas dire. Principes de Semantique Linguistique. Collection Savoir.
- Harris, Zellig S. (1951; 1960), Structural Linguistics. Phoenix.
- Grice, H. Paul (1975), "Logic and conversation," in Peter Cole and Jerry Morgan, eds., Syntax and Semantics Vol. III: Speech Acts. Academic Press. (pp. 41-58)
- Hockett, Charles F. (1958), A Course in Modern Linguistics. Macmillan.

Jakobson, Roman (1936; 1966), "Beitrag zur allgemeinen Kasuslehre," TCLP 6.240-288; reprinted in Eric Hamp, Fred Householder and Robert Austerlitz, eds., Readings in Linguistics II. University of Chicago Press. (pp. 51-89)

Katz, Jerrold J. (1972) Semantic Theory. Harper and Row.

Katz, Jerrold J. and Jerry A. Fodor (1963), "The structure of a semantic theory," Language 39.170-210.

Kiparsky, Paul (1976), "Oral poetry: some linguistic and typological considerations." in B. A. Stolz and R. S. Shannon, eds., Oral Literature and the Formula. University of Michigan.

Leech, Geoffrey (1974), Semantics. Penguin.

Levi, Judith (1978), The Syntax and Semantics of Complex Nominals. Academic Press.

Makkai, Adam (1972), Idiom Structure in English. Mouton.

Morgan, Jerry L. (1978), "Two types of convention in indirect speech acts," in Peter Cole, ed., Syntax and Semantics Vol. IX: Pragmatics. Academic Press.

Nunberg, Geoffrey (1977), "The Pragmatics of Reference" C.U.N.Y. Ph.D. Dissertation.

Searle, John (1978), "Literal meaning," Erkenntnis 13.207-224.