Are There Semantic Natural Kinds of Words?

KENT JOHNSON

Abstract: Gareth Evans proposes that there are semantic natural kinds of words. In his development of this theory, he argues for two constraints on the identification of these kinds. I argue that neither of these constraints are justified. Furthermore, my argument against Evans' second constraint constitutes a direct argument for the existence of semantic natural kinds, something Evans himself never offers. I conclude by sketching some positive details of a more plausible theory of semantic natural kinds.

Introduction

Consider the following inferences:

(1) Molly is a Republican chemist; therefore Molly is a chemist.
(2) There are three apples in the basket; therefore there is a prime number of apples in the basket.

Unlike (2), the inference in (1) is normally thought to be due to its semantic structure. Barring irrelevant subtleties, the semantic structure of (1) might be thought to be *Republican(Molly) & chemist(Molly)*; *therefore, chemist(Molly)*. Now consider the inference:

(3) John is a large man; therefore, John is a man.

The first sentence in (3) does not have the structure *large(John) & Man(John)*. As is well-known, *large* is an attributive adjective. To say that John is a large man is to say that John is large *for a man*; he may not be large for a basketball player. So we need to make the semantic properties of *large* sensitive to the noun it modifies. This kind of sensitivity is found in other adjectives, such as *fake*, which appear in the *invalid* inference in (4):

(4) X is a fake gun; therefore X is a gun.

Special thanks to Ernie Lepore for reading many drafts of this paper and providing a great deal of criticism. I would also like to thank Elisabeth Camp, John Carroll, Robert Matthews, and Stephen Neale, and two anonymous referees for *Mind & Language* for their useful feedback. A shorter version of this paper was given at the 2001 Pacific APA meeting; I am grateful for the comments I received there.

Address for correspondence: Department of Logic and Philosophy of Science, 3151 Social Science Plaza A, University of California, Irvine, Irvine, CA 92697, USA.
Email: johnsonk@uci.edu

Mind & Language, Vol. 18 No. 2 April 2003, pp. 175–193.

Here we see that *fake* too must be sensitive to the nominal, since X might be a fake gun, but not a fake garden-hose nozzle. Is there any way that we could represent the inference in (3) within a semantic theory without also predicting the inference in (4)? More importantly, is there any reason that we *should* represent the inference in (3) as being due to its semantic structure? Could (3) be more like (2) than (1) in that its validity is not due to its structure? In 'Semantic Structure and Logical Form', Gareth Evans argues that these questions admit of principled answers if we treat words as falling into semantic natural kinds (or as he calls them, semantic categories). According to Evans, semantic categories specify:

> for each kind of semantic expression, an entity—a set, a truth value, a function from sets to truth values, or whatever—which may appropriately be assigned to members of that *kind* upon an arbitrary interpretation of the language. We can regard the specification of the kind of assignment as a specification of the underlying real essence which a word has in common with many other words, and of which the validity of certain inferences involving it is a consequence. These will be the structurally valid inferences; inferences which are truth-preserving no matter how we permute assignments *within the limits laid down as appropriate for members of that category* (Evans 1976, p. 61 [emphases added]).

From a formal perspective, Evans argues, whether or not (3) is explained by the relevant semantic categories amounts to whether the word *large* belongs to a category C that is specified as something like (5) or as something more like (6).

(5) a. For any term ϕ in category C: x satisfies 'ϕ N' iff x is ϕ with respect to the property of being an N;
    b. [So: x satisfies *large man* iff x is large with respect to the property of being a man]

(6) a. For any term ϕ in category C: x satisfies 'ϕ N' iff x is ϕ with respect to the property of being an N, and x is an N;
    b. [So: x satisfies *large man* iff x is large with respect to the property of being a man and x is a man] (cf. Evans 1976, p. 56)

(6) but not (5) builds into the meaning of *large* that a large N is an N.

The philosophical issues about which semantic category a word belongs in are not formal problems. As philosophers, we face two questions: (i) Are there semantic categories of words in natural languages like English?, and (ii) If such categories exist in a language, can we know enough about them to use them in linguistic explanations? Both questions are important. If categories do not exist, then a positive answer to the second question only shows that they are useful for languages other than the ones we speak. On the other hand, securing the existence of semantic categories is interesting only if we can discern enough of their properties to make use of them. Concerning question (i), Evans offers no help. He simply

assumes that semantic categories exist, and explores question (ii). In fact, given that he employs semantic categories to determine whether inferences like (3) are due to semantic structure, his interest in question (ii) is restricted to a narrow class of linguistic phenomena. In an attempt to show that semantic categories can provide principled answers to questions about inferences and semantic structure, Evans offers two constraints on the identification of semantic categories that do not rely on the data from any particular theory. Although I am sympathetic to linguistic natural kinds, I don't think that either of Evans' constraints are justified. In the next two sections, I will discuss and argue against each of these constraints in turn. My argument against the second constraint will also supply some evidence for the existence of semantic categories (although not the ones Evans thinks exist). Thus, my rejection of his second constraint will supply a missing argument for a positive answer to question (i). In the final section, I will outline a theory of what these other semantic categories might be like.

Before beginning, a word on interpretation is in order. I maintain that my interpretation of Evans is correct, and I will take some pains to defend it. However, a chief goal of this paper is to explore two interesting methodological principles about the nature of language and the construction of semantic theories. Regardless of whether Evans held these principles (I will argue that he did), we will see that they are sufficiently interesting in their own right to merit some attention.

1. First Principle: The Maximizing Constraint

How could we identify semantic categories? We have already seen that Evans associates each semantic category with 'an entity—a set, a truth value, a function from sets to truth values, or whatever' (p. 61). This entity is intended to be the 'real essence' from which the inferential properties distinctive of the category 'flow' (*ibid.*).[1] Evans' idea here is that this logical entity will specify those inferential properties that a word in the associated semantic category has because the word is a member of that category. Is there anything more we can say about the nature of the logical entity that serves to individuate a given semantic category? There is: there are a couple reasons why it is natural to identify the logical entity associated with a given semantic category with the inferential properties that the entity

[1] Although I used a sentential form in (5a) and (6a), these could easily be replaced with functions, in the spirit of e.g., Montague 1974. Using such structures, then, the logical entities corresponding to (5a) and (6a) would be:

(5a′) $\lambda N \lambda x[\phi(x, \hat{}N)]$
(6a′) $\lambda N \lambda x[\phi(x, \hat{}N) \,\&\, Nx]$

(The carat is Montague's abstraction operator, which takes us from a predicate like the meaning of a noun to the property that that predicate expresses. In this way, a predicate can fill an argument position in ϕ.)

specifies. In the first place, the primary purpose of positing semantic categories at all was to determine which inferences are due to the structure of the language and which are not. So by identifying the logical entity with the inferential properties it specifies, we are merely saying that the logical entity is that which determines the inferential properties associated with a given category. In other words, the logical entity is that which accomplishes the task at hand. In the second place, if we did not identify a word's semantic essence with its inferential properties, we would have to identify this essence with some other thing E. But it is entirely unclear what E could be, and it is unclear what it would be for the inferential properties to 'flow' from E. In any case, one cannot just stipulate that words have semantic essences of a certain type; one must also make a plausible case for the existence of such essences. Evans offers no account of what such essences might be. In the absence of reasons for thinking that words might have semantic essences above and beyond their inferential properties, I will make the following assumptions: (i) two semantic categories are identical if and only if they are associated with the same logical entity; (ii) two logical entities are identical if and only if they specify the same set of inferential properties, and (iii) a word belongs to a given semantic category only if the word has all the logical properties that are associated with that category.

(As we will see, Evans' two principles for individuating semantic categories conflicts with the 'extensional' assumption (given by (i), (ii), and (iii)) that a semantic category can be identified by the set of inferences associated with it. This conflict points the way towards an independently motivated positive theory of linguistic natural kinds. Although some inferential properties 'flow' from these natural kinds, they satisfy neither of Evans constraints.)

So far, Evans' theory has given us a criterion for the *individuation* of semantic categories. However, we do not yet have a theory for the *identification* of any particular category. In order to determine the appropriate category for *large*, for instance, we still need to determine which inferential properties are shared by all the members of the category. To do this we need to determine whether e.g., *fake* is an element of the same category as *large*. If *fake* and *large* are in the same category, then the category cannot license the inference Adj N $\Rightarrow$ N, where N is a nominal and Adj is an expression in the category. It is at this point that Evans' principles for semantic category indivduation are designed to be of help. Evans' first principle appears when he writes:

> *What we expect, then, is the provision of the most determinate and yet economical statement of the kind of semantic contribution made by any expression of a given type, thereby making structurally valid as many inferences as possible.* . . . [W]e aim at the sort of illumination that can come from an economical axiomatization of the behaviour of groups of expressions. Then we can say: '*This* is the kind of expression *e* is, and that is why these inferences are valid.' (p. 63 first emphasis added).

Evans' idea here is to seek out a theory of semantic categories that represents the most inferences for the least overall cost in terms of the number of semantic

categories posited. So when constructing a theory of semantic categories, you try to keep it simple, resisting added complexity. At the same time, you also try to maximize the number of (commonly known) inferences that are explained by the semantic structure of the theory. Evans suggests these two desiderata will keep each other in check: by maximizing the theory's simplicity, you resist adding semantic categories. And maximizing the number of inferences explained means increasing the number of categories, because as words are split into more and smaller categories, the members of each category will have more inferences in common. The hope is that by aiming at a maximal joint satisfaction of these two desiderata, the construction of a semantic theory will achieve a certain degree of uniqueness. If so, then one would have some right to claim that the inferences the theory declares to be structural really are structural. Let's call the above principle for individuating semantic categories the *maximizing constraint*, summarized as follows:

(7) *The Maximizing Constraint*: Semantic categories can be identified as those that are posited by the simplest correct semantic theory that represents the most inferences.

(Here a 'correct' semantic theory is one that meets every desideratum on a semantic theory other than satisfying the Maximizing Constraint. Since one of the duties of a theory of semantic categories is to help determine whether certain inferences are due to the semantic properties of words, capturing or not capturing one of these inferences will not be an (antecedently identified) desideratum on any semantic theory of the language.)

To evaluate the Maximizing Constraint, let's make the simplifying assumptions that we know how to measure the simplicity of a theory, and that in the present case this measurement amounts to determining how many semantic categories the theories each employ. Let's also assume that we have some trouble-free way to identify inferences and inferential properties.[2] I want to argue that if some valid inferences are not explained by the semantic structure of the component sentences, then the maximizing constraint is unjustified. The assumption that there are valid

---

[2] These latter assumptions are nontrivial. An appropriate method for individuating inferential properties keeps us from collapsing the number of semantic categories and yet retaining lots of semantic inferences by allowing such inferential properties as

(★) If A is an intersective adjective and N is a noun, then anything that satisfies $[_{N'}$ A N] also satisfies N.

(An intersective adjective is one that modifies a nominal only by reducing the extension of the nominal in some way.) Inferential properties like (★) would render the whole project trivial, so we should assume that they are ruled out. It is also important to know how to enumerate the inferences. We want to capture the 'most' inferences using the smallest number of semantic categories. But such inferences as *large* $N' \Rightarrow N'$ comprise a countable infinity of particular inferences (e.g., *large basketball player* ⇒ *basketball player, large survivor of at least three forms of cancer* ⇒ *survivor of at least three forms of cancer*).

inferences that are not explained by a semantic theory is especially plausible if we follow Evans in assuming that the purpose of a semantic theory is to tell us what we know (perhaps tacitly) in virtue of which we know the meanings of the expressions of a language (e.g., 1976, p. 51, 1981, p. 328). According to this view, the job of a semantic theory is to teach us something about the mind (or human behavior; cf. Evans, 1981) by describing what semantic structure it assigns to a given syntactic form. It doesn't follow on this view that we should represent an inference as structural simply because we can. Even if we all know that *seven flowers bloomed* implies *a prime number of flowers bloomed*, this knowledge may be due to our cognitive mathematical competence and not to our knowledge of the meaning of the sentences. So I will assume that some valid inferences are not explained by our knowledge of language.

Having made these simplifying assumptions, let's see how the Maximizing Constraint works. Suppose we are constructing a meaning theory and are left with a set X of leftover inferences, and we don't know whether to count them as due to their semantic structure.[3] The maximizing constraint tells us to make the inferences semantic if it doesn't cost too much in terms of the number of semantic categories posited. But there is no reason to be biased towards construing inferences as due to semantic structure. Since we're assuming some inferences are not explained by their semantic structure, they must be explained by some other theoretical resources. Why then shouldn't we be biased towards treating the inferences in X as explained—at least in part—by these other resources (unless the theory constructed so far gives us reason to do otherwise)? There does not appear to be any methodological reason for representing these inferences within the semantic theory, as opposed to somewhere else. There is however a methodological constraint that is relevant. The constraint is that all the inferences in X should be treated the same way, except for those that we later find reasons for treating differently. To see this, remember that by hypothesis we have no evidence

[3] We can suppose that there are numerous inferences in X which we have independent reasons for treating as semantic or as non-semantic. Clearly there are inferences which are not part of our semantic knowledge, structural or otherwise. Any theory that represents me as knowing (2) in virtue of my knowledge of English is surely false. Similarly, as Evans himself noted (Evans, 1981), there can be empirical evidence for the semantic structure a theory assigns to certain expressions. In psycholinguistics, such evidence takes many forms: patterns of first language acquisition, patterns of language deficits resulting from head injuries, errors made when processing under suboptimal circumstances, eye movements when reading, and various sorts of measurements of brain activity. In linguistics, further evidence comes from regularities both within a single language and across various languages. We should interpret the maximizing constraint as applying to the remaining inferences for which there is no evidence one way or the other. (It's worth noting a couple more powerful assumptions Evans needs. First, we must assume that after we have carved away those inferences whose semantic status is independently established, there are still enough inferences left that it matters which of them we count as semantic. Secondly we need to assume that once we've dealt with the independently established inferences, we are not left with such a strange and arbitrary organization of words (inferentially speaking) that there is no chance that the positing of semantic categories could do some theoretical work.)

for treating the inferences in X one way or the other. If we explain the inferences in X using *both* the resources of a theory of semantic structure *and* the other resources for explaining nonsemantic inferences, we have granted ourselves a large explanatory apparatus. If instead we try to explain the inferences in X using *only* the semantic theory or *only* the other resources, but not both, then either way we will be explaining the phenomena using a less powerful explanatory apparatus. *Ceteris paribus*, if a more restrictive methodology can produce an explanatory theory, then that methodology should be preferred over a less restrictive one. Thus, we should try to treat all the inferences in X as semantic *or* as all non-semantic. Seen from this light, the maximizing constraint simply declares that we should adopt the first disjunct. But without an argument why this should be so, we don't really have any reason to adopt it over a rival principle that declares that the inferences in X should be assumed to be non-semantic. Thus, the maximizing constraint begs a crucial methodological question. Without any additional support for the answer it assumes, the principle is unjustified.[4]

The methodological argument has some potentially far-reaching consequences, which I will only mention here. The point made entails that one cannot simply assume without argument that a given inference is due to its semantic structure. It could turn out that many commonly known inferences are best treated as non-semantic. The inference $P \Rightarrow P \vee 17 + 39 = 56$ strikes me as a good example. Another example: personally, I do not find it obvious that the semantic structure of *Tanya persuaded Lana to leave* implies *Lana intended to leave*, although Chomsky

[4] An anonymous reviewer offered two useful comments about the argument just given. In the first place, depending on the nature of the particular semantic theory (and the goals in constructing such a theory), the simplicity in question might not be restricted to simply the number of semantic categories. Instead, their simplicity might also depend on some other factor, such as the 'complexity' of the properties used in the individuation of the semantic categories. While this is certainly true, it is irrelevant to the above argument. Regardless of how you measure simplicity, we can still ask why the winning theory is the correct one that fares best along both dimensions of simplicity and maximizing the number of inferences in X that are represented within the semantic theory. (In other words, the present issue is not concerned with how simplicity is measured, but with the justificatory role of the number of inferences in X that the theory represents.)

The second comment is that one might simply stipulate that there is a default bias in favor of a semantic theory explaining inferences. So in the situation imagined above, where we are at a theoretical impasse, it would be methodologically appropriate to treat the inferences in question as due to their semantic structure. For all I know, if one has the right goals in constructing a theory of language, such an objection is well placed. However, if one's goal is the construction of a psychologically plausible characterization of the structure of a speaker's knowledge of language, it is unclear why there should be any such default bias. We know that there are some commonly known inferences that are due to rational abilities that are best thought of as distinct from our linguistic abilities. So to impose a default bias on this sort of linguistic project would be to stipulate that, all other things being equal, an inference is due to our linguistic abilities, not to our general abilities to draw inferences. I fail to see what could make such a default bias plausible. (Indeed, if there was any default bias here at all, which I doubt there is, I would think that it would be in favor of the inferences being due to our general abilities to infer, rather than to their being artifacts of the mechanism (i.e. language) by which we express these inferences.)

assumes it does (Chomsky, 1995, p. 17; Chomsky, 2000, pp. 62, 176–77). I can agree that one must know the above inference in order to be counted as a rational speaker (and perhaps for us to be able to discern that she is a speaker at all). But the present issue is not about the set of inferences a rational speaker knows. It is about a proper subset of those inferences, namely the ones that speakers know in virtue of knowing the language. (Nobody doubts that the inference *Large* $N' \Rightarrow N'$ is in the former set; the issue is whether it is in the latter set as well.) In a related vein, whether or not an inference is semantic should also not depend on the standard first-order (or higher-order) representation of these sentences. Rather, the issue is whether the English sentences should be assigned logical structures in some *appropriate* logic or other formal system so that the structure assigned to the second sentence is a logical consequence of the structure assigned to the first (using an appropriate notion of logical consequence). What is to count as the correct logic and the correct notion of logical consequence are not trivial questions (cf. Montague, 1970 for some general discussion). There is an even deeper question of what could motivate the addition of structue to a semantic theory for the sake of capturing an inference. A theory of semantic categories might be able to partly address this question, if such a theory itself can be defended. (Higginbotham has made some additional proposals that concern the facts a semantic theory should explain; e.g., Higginbotham, 1992, 1988, pp. 237–9, 1989, esp. pp. 465–71, 1989b.) In short, the issues surrounding the justification of the maximizing constraint quickly lead to some of the most fundamental issues in the philosophy of language. Further discussion of this matter is a topic for another paper. I will therefore leave the matter where it stands and turn to Evans' second principle.

2. Second Principle: The Different Explanation Constraint

Evans himself was not fully convinced that the maximizing constraint would be enough to identify the semantic categories, although not for the reasons I have given (cf. p. 63). So he argues for an additional constraint on the identification of semantic categories. He writes, 'We should regard our construction of, and assignments to, categories in the following spirit: if two expressions behave in the same way but are in different categories, this is a lost generalization' (p. 64). From this he infers the methodological principle:

(9) *The Different Explanation Constraint*: Any two distinct categories must provide 'a *different* explanation for the behaviour which members of the [first] category [have] in common with the [second]' (*ibid.*).

'By imposing this requirement', he concludes, 'we make the notion of structural validity transcendent [sc. theory-neutral]' (Ibid.). So if two distinct categories both explain a similar inference that their respective elements participate in, then they must explain that inference as holding for their elements for different reasons. *Prima facie*, the different explanation (DE) constraint seems reasonable. Evans' point in

introducing semantic categories was to arrange words into natural kinds based on some shared inferential properties. He'd like to be able to say that a given word has some semantic feature that gives it certain inferential properties *because* it is a member of a certain category. Since semantic categories are supposed to explain certain kinds of inferences, if a theory says that the members of various categories participate in the same inference for the same reason, then that is evidence that the theory has not identified the categories correctly.

Aside from worries about how much the DE constraint actually contributes to semantic theories, the real problems with it are that it is question-begging and false. It is question-begging because it excludes two semantic categories from being subcategories of a larger one which itself serves to explain some of the inferential behavior of the two contained categories. (We can define: category X is a *subcategory* of Y iff all the inferential properties associated with X are also associated with Y, and all the words in X are also in Y.) To see this, imagine that there were semantic categories A and B that were subcategories of a larger category C. Now imagine that C was defined by its explaining a certain inference *i* that all its members participate in. Thus, all the words in A and B participate in *i*, and for exactly the same reason, namely because all these words are members of C. But now the DE principle declares that the identification of the categories A, B, and C must be modified to prevent this result.[5] So the DE principle rules out semantic subcategories. Even more generally, it rules out the possibility that semantic categories could partially overlap. The fact that the DE constraint rules out such possibilities begs an interesting and important question about the nature of human languages. So regardless of what natural languages are actually like, unless we know that they do not contain overlapping semantic categories, the DE constraint is anything but theory neutral. What is worse, in addition to being question-begging, natural languages seem to be organized in such a way that the DE constraint makes a false generalization about them. There are a number of plausible examples that contradict the DE principle, but I will focus on just one. Of course, the example at best shows that the DE constraint is probably false. Individual words don't wear their semantic categories on their sleeves (or in their morphosyntactic form, for that matter), so one could always maintain that the DE constraint is correct, and simply manipulate the semantic categories to suit the data. My present goal is only to argue that the price of doing so is too high to merit salvaging the DE constraint.

In the last twenty years or so, one the most heavily studied types of inferential phenomena of natural language has been the monotonicity effects of quantifiers (e.g., Barwise and Cooper, 1981; Higginbotham and May, 1981; Keenan and Stavi 1986). In the frameworks in which these phenomena are typically studied, quantifiers are thought of as relations between predicates (i.e., between expressions that can be true or false of an entity). Thus, *All As are Bs* is true iff $\mathbf{A} \subseteq \mathbf{B}$, where '**A**'

[5] That the DE principle should be interpreted so that no two categories should be allowed to share any common inferences is clear from the text (cf. p. 64).

and '**B**' denote the extensions of A and B. Similarly, *Some As are Bs* is true iff $\mathbf{A} \cap \mathbf{B} \neq \emptyset$, *Most As are Bs* iff $|\mathbf{A} - \mathbf{B}| < |\mathbf{A} \cap \mathbf{B}|$, and *More As than Bs are Cs* iff $|\mathbf{B} \cap \mathbf{C}| < |\mathbf{A} \cap \mathbf{C}|$. Viewing quantifiers as relations between predicates allows us to view predicates as arguments. In this sense, we can define: A quantifier Q is *monotone increasing* in its nth argument iff whenever $\mathbf{A}_n \subseteq \mathbf{B}_n$, $QA_1 \ldots A_n \ldots A_m$ implies $QA_1 \ldots B_n \ldots A_m$. Similarly, a quantifier Q is *monotone decreasing* in its nth argument iff whenever $\mathbf{B}_n \subseteq \mathbf{A}_n$, $QA_1 \ldots A_n \ldots A_m$ implies $QA_1 \ldots B_n \ldots A_m$. We can indicate facts about monotonicity by replacing arguments of a quantifier with '↑', '↓' and '−' to indicate that the quantifier is monotone increasing, decreasing, or neither in that argument place. With that said, we can observe the small sampling of facts in (10), with data supporting them in the table below.

(10) a. Every ↓ ↑
b. No ↓ ↓
c. Some ↑ ↑
d. Most − ↓
e. More ↑ ↓ −

a. Every dog was tired ⊨ Every brown dog was tired. Every brown dog was tired ⊭ Every dog was tired.	Every dog was tired and hungry ⊨ Every dog was tired. Every dog was tired ⊭ Every dog was tired and hungry.
b. No dog was tired ⊨ No brown dog was tired. No brown dog was tired ⊭ No dog was tired.	No dog was tired and hungry ⊭ No dog was tired. No dog was tired ⊨ No dog was tired and hungry.
c. Some dog was tired ⊭ Some brown dog was tired. Some brown dog was tired ⊨ Some dog was tired.	Some dog was tired and hungry ⊨ Some dog was tired. Some dog was tired ⊭ Some dog was tired and hungry.
d. Most dogs were tired ⊭ Most brown dogs were tired. Most brown dogs were tired ⊭ Most dogs were tired.	Most dogs were tired and hungry ⊨ Most dogs were tired. Most dogs were tired ⊭ Most dogs were tired and hungry.
e. More dogs than cats were tired ⊭ More brown dogs than cats were tired. More brown dogs than cats were tired ⊨ More dogs than cats were tired. More dogs than cats were tired ⊨ More dogs than black cats were tired. More dogs than black cats were tired ⊭ More dogs than cats were tired.	More dogs than cats were tired ⊭ More dogs than cats were tired and hungry. More dogs than cats were tired and hungry ⊭ More dogs than cats were tired.

The quantifiers in (10) have different monotonicity properties. According to Evans' theory, each quantifier belongs to exactly one semantic category. Now, if the facts about monotonicity are structural, then we will need 'a *different* explanation' for the fact that *every* and *no* are monotone decreasing in their first argument. This creates a real problem when we turn to see how the theory would interact with some well-known generalizations in linguistics. One such generalization concerns the so-called 'negative polarity items', expressions such as *any, ever, yet, all that, give a damn*, etc. A fact about them is that they are restricted to certain kinds of semantic environments. With respect to quantifiers, the generalization is:

(11) Negative polarity items can appear in the argument of a quantifier if that argument is monotone decreasing.

Without any other 'negative' element to license them, negative polarity items are unacceptable.[6] For *no, some* and *every*, (11) is exemplified in (12):

(12) a. No [men with *any* sense] [cheered];
b. No [animals] [were *ever* hurt during the test];
c. *Some [women with *any* sense] [cheered];
d. *Some [animals] [were *ever* hurt during the test];
e. Every [worker who *gives a damn*] [will keep his mouth shut];
f. *Every [child] [gave *any* money to charity];

The problem with all this is that (11) cannot be stated if we adopt the DE constraint (unless there is additional apparatus present). How could it? If quantifiers like *some, no* and *every* are placed in the *same* semantic category, then that category will not be able to acknowledge the monotonicity properties of these words. The category could not acknowledge that *no* and *every* are monotone decreasing in their first argument because of the presence of *some* in the category; similarly, the presence of *some* and *every* prevents the category from acknowledging that *no* is monotone decreasing in its second argument. If the category cannot explain these properties, then the category will be of no use in explaining (11). On the other hand, if *some, no* and *every* are placed in *different* categories, those categories *still* cannot explain (11), because the DE constraint demands that there be different reasons why the two quantifiers license the inferences that they do. So unless we add some further apparatus to explain how these categories came to be, there can be no saving (11) within the theory, and this shared ability to license negative polarity items remains a mystery. Losing such a generalization is a bad thing, since the desire not to lose them is precisely what motivated the DE constraint, which started this whole business.

---

[6] The generalization can be found in many standard textbooks, e.g., Larson and Segal, 1995, pp. 281–286; Chierchia and McConnell-Ginet, 1990, pp. 421–425; cf. also Ladusaw 1996. This generalization has not gone unchallenged: cf. Linebarger, 1987 and Ludlow, forthcoming.

What went wrong with Evans' theory? The problem is that by adopting the DE constraint Evans' theory does not allow for semantic categories to overlap. If semantic categories could overlap, the DE constraint would be trivial: whenever you want two expressions to be from different categories and yet to underwrite the same inference for (what appears to be) the same reason, just make that inference distinctive of a more general category to which the two expressions belong. The two expressions will still, on the present story, belong to different categories, because they will not belong to all the same general categories. (In this way there can be as many categories of expressions as there are subsets of inferences used to individuate categories.) Even if we only allowed semantic categories to overlap some of the time, the DE constraint would still be ineffectual as a methodological principle. In such a case, it would amount to the injunction: 'Different categories shall provide different explanations for similar inferences, except when they don't'. Clearly, such an injunction does not help us to build a semantic theory.[7]

Reminiscent of the beginning of Section 1, a defender of Evans' theory may wish to object to the present argument by claiming that the various quantifiers are in different categories. The quantifiers are in different categories, the objection goes, because each quantifier has some further essence from which its monotonicity properties 'flow'. In reply to this, we can note that until we have a plausible theory of semantic essences from which monotonicity properties 'flow', the objection is not yet fully-formed. I know of no theories of semantic essences which could both place the quantifiers we have looked at in different semantic categories, and which would somehow explain or entail the fact that the quantifiers have the monotonicity properties that they do. (Moreover, outside of this discussion of semantic categories, it is unclear why anyone would want or expect quantifiers not to be cross-classified in the way that they appear to be. In general, cross-classification, is quite common; the partitioning of words into disjoint subsets is not especially common in linguistics.)

In short, the DE constraint prohibits semantic categories from overlapping. But the empirical facts seem to suggest that inferential patterns overlap one another rather than clustering neatly together. So it looks like the DE constraint begs a crucial question about natural languages, and in doing so it assumes an incorrect answer. Of course, one could preserve the DE constraint by simply denying that the monotonicity properties of quantifiers are structural. But then one would like to know whether anything motivates such a claim other than the desire to preserve the DE constraint.

On a more positive note, the facts about licencing negative polarity items provide a kind of evidence for the semantic categorization of quantifiers that does not rely on either of Evans' constraints. The facts in question conflict with the DE constraint, so they trivially do not rely on it. As for the maximizing constraint, its purpose is to motivate some way of dealing with those inferences which we don't have enough reason to treat as semantic or as non-semantic.

---

[7] Notice, incidentally, that such an injunction would not even motivate a default bias regarding the seeking of explanations of inferences.

However, we have just seen that the linguistic mechanisms need to be sensitive to the monotonicity properties of quantifiers, and monotonicity properties are inferential properties. Thus, there is no reason to appeal to the maximizing constraint, because we already know that these inferential properties are due to semantic structure. Thus, in rejecting the DE constraint, we have also generated some independent evidence for the existence of (overlapping) semantic categories.

Let me summarize the discussion so far. We began with a general question about what kind of structure should be represented within a semantic theory. The need for a theory of semantic structure appears when we ask whether a given inference is due to its linguistic properties alone. Evans suggests that these questions can be answered by finding the right theory of semantic natural kinds (i.e., categories). Since these categories associate inferential properties with sets of words, we need some principled way to pick out a correct theory of semantic categories from all the logically possible alternatives. Evans' first principle for doing this was the maximizing constraint. But this principle simply declared that we should try to explain valid inferences in the semantic theory, and there's no reason that we should be biased towards supposing that our knowledge of language explains our knowledge of a given valid inference. Evans' second attempt was the DE constraint. This constraint told us that distinct categories may not licence the same kind of inference for the same reason. But this constraint begs a crucial question about the possibility of overlapping semantic categories, and in fact, it seems that quantifiers fall into a group of overlapping categories.[8] So Evans has provided no reason to favor one theory of semantic categories over another. Nonetheless, I think that the notion of semantic natural kinds of words may be theoretically important. In the next section, I will try to sketch this view and defend its importance.

3. A Different Theory of Semantic Kinds

In order to see why we might want to posit semantic natural kinds, it will help to step back for a moment to think about what a semantic theory is for. A common view often associated with Chomsky is that a semantic theory is part of a larger theory of the cognitive ability normal humans have to become able to speak and understand the language(s) they are exposed to at a young age (for details of Chomsky's own (current) views, see the papers in Chomsky, 2000). The job of the semantic theory is to state the meanings of the expressions (words, sentences, etc.) of the language. Crucially, however, there is no guarantee that this semantic theory will be much like the formal semantics developed for artificial logical languages. In standard formal languages, the

[8] Ernie Lepore has pointed out to me that the problem of overlapping categories could arise even for the attributive adjectives Evans considers. After all, if x is a large N, then x is large, and if x is a fake N, then x is (a) fake. If *large* and *fake* were placed in distinct categories, and the above inferences were treated as structural, the problem might arise. However, a clear conception of the semantics of *large* and *fake* when not modifying a N′ would also be required.

syntax of the language is generated independently of the semantics, and semantic properties are afterwards associated with expressions. However, there is evidence that a true theory of the mind will posit more interaction between the syntax and semantics of natural language. In particular, there is evidence that many simple-looking words are actually highly semantically structured, and this semantic structure serves to explain some of their *syntactic* properties. Let me illustrate this with an example. Consider the verbs *spray*, *cover*, and *throw*. These verbs appear in distinct but overlapping syntactic environments:

(13) a. Sarah *sprayed* the windows with cold water;
a′. Sarah *sprayed* cold water onto the windows;
b. Sarah *covered* the table with the dishes;
b′. *Sarah *covered* the dishes onto the table.
c. *Sarah *threw* the table with the dishes;
c′. Sarah *threw* the dishes onto the table;

The patterns in (13) are not isolated. Many verbs pattern like *spray*: e.g., *load, cram, plant, smear, splash, encrust, strew, wrap, empty*. Other verbs pattern like *cover*: e.g., *decorate, fill, soil, bathe, saturate, infect, taint, pollute, litter, ornament, pave*. Finally, many verbs behave like *throw*: e.g., *pour, arrange, immerse, lodge, mount, pour, spill, coil, shove, funnel*. (Levin, 1993, pp. 50–1, and Pinker, 1989, pp. 126–30 contain many more verbs in all three classifications.) What could account for this distribution of verbs? A careful look at the three classes of verbs given in (13) shows that they have a certain semantic similarity. *spray* seems to express something about both what happened to the windows and how the water was moved onto them. *Cover* on the other hand, only describes what happened to the table. *Throw* however, seems to only describe how the dishes were moved. As more data like this piles up, patterns emerge that relate the kind of event (or situation or relation or whatever) that the verb denotes with the kinds of syntactic patterns it can appear in. A theory can acknowledge these patterns only if it can organize words into semantic categories. As our small example suggests, these categories may overlap in multiple ways. A common way of identifying them is by means of a severely restricted language composed of a small number of predicates. Ignoring the subtle details that keep lexical semantics from being a completed field, we can say the syntactically relevant semantic properties of *spray*, *cover* and *decorate* can be given as:

(14) a. *Alternating Verbs*: Subject changes Place by causing Stuff to go into/onto Place in a certain way.
b. *With-Only Verbs*: Subject changes Place.
c. *Into-Only Verbs*: Subject causes Stuff to go into/onto Place in a certain way.

The semantic explanations of these verb classes is robust. As new verbs in these semantic classes enter the language, they obey the grammatical patterns predicted by them. For instance, you might report that you downloaded a file onto your hard

drive, but you wouldn't say that you **downloaded your hard drive with a file*, since the verb *download* does not require that the site of downloading be changed (or 'completely affected'; cf. Tenny, 1994) by the downloading. Similarly, you might ftp a file into your email account, but you would not **ftp your email account with a file*.

Roughly speaking, semantic categories can be identified with these simple formulae. However, why should we believe that these formulae are actually part of the language? More to the point, why should we believe that the basic predicates express concepts that are fundamental to language? Although they neatly capture the distributions of the verbs in (13), they seem to compose a pretty heavy-duty piece of apparatus. If all they do is account for distributions like (13), then it is not clear that the work they do outweighs the price of positing them. Fortunately, though, the concepts appealed to in (14) do much more than account for locative verb distributions. I will briefly mention four other areas where they appear and do work. First of all, they explain numerous other grammatical distributions and semantic properties. For example, the concept CAUSE appears not only in the alternations in (13), but it is also essential to those intransitive verbs that have causative counterparts, as in *the water boiled, John boiled the water*, and *the sandcastle toppled, Betina toppled the sandcastle*. It also appears in resultative clauses, such as *Marsha shot Susan dead*, where this means that Marsha shot Susan with the result that Susan became dead. Secondly, the concepts used in (14) appear to structure other languages as well. A glance at the literature shows these concepts serving explanatory functions (often quite similar to the ones they serve for English) in such diverse languages as Japanese, Berber, Chukchee, Icelandic, Turkish, Russian, and Edo (e.g., Comrie, 1985; Baker, 1988; Tenny, 1994; Bittner, 1999).[9] Thirdly, these concepts figure into explanations of child language acquisition and adult language processing (e.g., Pinker, 1989; Gropen, Pinker, Hollander and Goldberg, 1991; Pinker, 1999; Tanenhuas *et al.*, 1993). In the case of locative verbs, for instance, Gropen *et al.* constructed scenarios where young children could use locative verbs. They found evidence that when a child uses a locative verb of the kind illustrated by (13b, b′) or (13c, c′) in the syntactic pattern that is ungrammatical in adult speech, the child also attaches an incorrect meaning to the verb. In fact, not only did the child appear to misinterpret the meaning of the verb, but she tended to assign it precisely the kind of meaning that would give it a semantic structure that would allow it to appear in the (for adults) ungrammatical form. Finally, sublexical concepts such as direct causation appear to be used by preverbal infants (e.g., Leslie, 1995, 1982; Leslie and Keeble, 1987). Leslie presented infants (some as young as 6 months) with different causal scenarios. The children displayed a differential reaction to the different scenarios, and this reaction tracked the distinction between direct causation and ordinary causation. These reasons make it

[9] An important caveat here is that the relation of causation may be somewhat less 'direct' when there is a causative morpheme present with the verb; however, the relation is not as unrestrictive as when the causation is expressed by a separate word, as in *cause to boil* (Comrie, 1985, p. 333; cf. Pinker, 1989, pp. 86, 95, 134).

plausible to suppose that the minds of speakers of natural language really do contain the sorts of concepts appealed to in the semantic categories described in (14). This claim is only an empirical hypothesis, subject to disconfirmation. Nonetheless, the evidence reviewed above provides strong support for it.

In actual linguistic practice, the sort of semantic structure I have been discussing is largely motivated by the grammatical patterns of various words and constructions. Studying such patterns is useful even in more complex cases than the ones I have looked at here. At the end of section 1, for instance, I said I did not find it obvious that *Tanya persuaded Lana to leave* implies *Lana intended to leave* by virtue of the structure of the two sentences alone. One might wonder, for instance, whether *persuade* could be analyzed as having a semantic structure of something like CAUSE TO INTEND. If this were the case, then the inference in question may indeed be due to semantic structure alone. However, determining that *persuade* has the semantic structure CAUSE TO INTEND is far from simple. For instance, if *persuade* does have a structure that contains CAUSE, then *ceteris paribus* we would expect *persuade* to behave like other verbs that contain this element. But *persuade* does not have the intransitive counterparts commonly found with causatives: we have *Tanya persuaded Lana to leave*, but we do not have **Lana persuaded to leave*; the judgments remain the same even when *to leave* is omitted.[10] These facts provide some initial evidence that *Tanya persuaded Lana to leave* does not contain a causative element (and so blocks one avenue for defending the structural validity of the inference in question). However, it may well be that the rather complex syntactic structure of the sentences is relevant, rendering the tests given here inconclusive (cf. e.g. Pesetsky, 1995).

Let me finish this outline by mentioning two philosophically relevant issues that arise within the study of lexical semantics. First, what is the precise nature of these concepts? For instance, the concept CAUSE appears in many places in lexical semantics, but it is clearly not identical with our concept of causation. Jill could cause the table to become loaded with dishes without loading it with dishes; perhaps she ordered one of the waiters to load it. However, the entailment in the other direction from *load the table* to *cause the table to be loaded* does appear to be valid. Such observations have led linguists to call the concept in question 'direct causation', since it is the existence of intuitively 'intermediary' causes that distinguishes the ordinary concept of causation from the linguistic concept. Under different guises and for different purposes, philosophers of action have already studied this concept extensively (e.g., Thomson, 1971, esp. p. 122; Thomson, 1970). For instance, the conditions for individuating actions and identifying the 'time of a killing' bear deep connections to direct causation. The second issue concerns what such theories tell us about the analytic/synthetic distinction. Instead of worrying about whether the inference *X is a bachelor, hence X is unmarried* is analytic, we might profit more by directing our

[10] Middle constructions are sometimes taken to be evidence of a causal element in semantic structure: cf. *the glass breaks easily* vs. **John resembles easily*. Here, too, *persuade* does not behave like a causative: **Lana persuades to leave easily*. I doubt that *?? Lana persuades easily* is grammatical, but I am unsure.

attention to verbs (cf. Moravcsik, 1981). We can ask whether e.g., (15a) analytically entails (15b). It may well turn out that if one understands these sentences then as a matter of psychological law, if one believes (15a) then one will believe (15b).

(15) a. Mary loaded the carts with boxes;
b. Mary loaded boxes onto the carts.

(Similarly, it may turn out that this inference meets other criteria for analyticity.) However, such an issue involves some nontrivial decisions about what information is semantically relevant. For instance, despite the apparent ease and obviousness of the inference in (15), when the two object NPs have definite determiners, the inference is no longer valid.

(16) a. Mary loaded the boxes onto the carts;
b. Mary loaded the carts with the boxes.

Neither sentence in (16) entails the other. Mary could load the boxes onto the carts even though many of the carts were left empty. Similarly, Mary could load the carts with the boxes, even though in the end she was not able to get all the boxes onto the carts.[11] (For more discussion of this phenomena, cf. the discussions of the so-called 'holism effect' of direct objects; e.g., Pinker, 1989, pp. 49–50, 67–69; Tenny, 1994, chapter 1.)

Despite the substantial differences between Evans' theory and the one just discussed, the two theories share several noteworthy affinities. For instance, they both agree that a semantic theory may contain more information about the denotation of a word than is needed to derive the meaning theorems of the sentences containing that word. In terms of deriving theorems that give the meaning of expressions of the object language, Evans' semantic categories are inert. From the perspective of the logic of the theory, they serve only to expand the number of inferences licensed by the theory. Similarly, the theory just presented treats many words as having semantic structure that is not required for generating the meanings of sentences. In fact, a semantic theory would be simpler if it simply gave the meanings of sentences like (13a, a′) using clauses in the spirit of (17):

(17) a. <x, e> satisfies [$_{V'}$ *load NP*] iff e is an event of x's loading the denotation of *NP*;
b. <x, e> satisfies [$_{VP}$ *V′ PP*] iff <x, e> satisfy *V′* and *PP*;
c. <x, e> satisfies [$_{PP}$ *with NP*] iff e is an event that occurs with the denotation of *NP*;
d. <x, e> satisfies [$_{PP}$ *onto NP*] iff e is an event that is onto the denotation of *NP*.[12]

---

[11] The failure of entailment in (16) can be even further strengthened by adding the particle *up*. Thus, although *I loaded up the boxes onto the carts, leaving the last two carts empty* remains acceptable, *I loaded up the carts with the boxes, leaving most of the carts empty* is clearly a contradiction.

[12] The use of events here is for illustrative purposes only. A function-argument structure along the lines of Montague, 1970 could generate meaning theorems just as trivially.

On the present theory of semantic categories, (17) cannot be the entire theory of the meaning of *load*. If it were, the linguistic theory would not have access to the semantic properties that drive its grammatical behavior.[13] When our assumptions about what a semantic theory is supposed to do are modified in the light of detailed empirical study, it may be that very many words will not receive the common and simple meaning axioms of the forms in (17). (For an extreme form of this last thesis, cf. Jackendoff, 1990.)

4. Conclusion

Evans' augmentation of a semantic theory with a theory of semantic natural kinds was wrong in its details. His two constraints on a theory of semantic categories are untenable. However, his idea that a semantic theory may be required to do more than articulate the meanings of the sentences of the language has received empirical support. By turning away from a word's inferential properties to its syntactic properties, we have seen that a theory of semantic natural kinds of words may very well have a place in a theory of language.

Department of Logic and Philosophy of Science
University of California, Irvine

References

Barwise, J., and Cooper, R. 1981: Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159–219.

Bittner, M. 1999: Concealed causatives. *Natural Language Semantics*, 7, 1–78.

---

[13] Similarly, the example of negative polarity item licensing and quantifier montonicity shows that a straightforward semantics as in (★★) cannot exhaust the semantic properties of quantifiers.

(★★) a. σ satisfies *Qx Ax are Bx* iff for Q functions σ' such that $\sigma' \sim_x \sigma$ and σ' satisfies *Ax*, σ' satisfies *Bx*.;
b. σ satisfies *All x Ax are Bx* iff for all functions σ' such that $\sigma' \sim_x \sigma$ and σ' satisfies *Ax*, σ' satisfies *Bx*.;
c. σ satisfies *The x Ax is Bx* iff for the function σ' such that $\sigma' \sim_x \sigma$ and σ' satisfies *Ax*, σ' satisfies *Bx*.;
d. σ satisfies *Many x Ax are Bx* iff for many functions σ' such that $\sigma' \sim_x \sigma$ and σ' satisfies *Ax*, σ' satisfies *Bx*.

(Here σ and σ' are metalinguistic variables ranging over functions that assign values to variables of the object language. The relation $\sigma' \sim_x \sigma$ holds iff σ and σ' differ at most in what they assign to the variable x.) If (★★) represents all that we know about the meanings of quantifiers, then there would be no explaining the fact that where npi's occur tracks the inferential properties of these words.

Chierchia, G., and McConnell-Ginet, S. 1990: *Meaning and Grammar*. Cambridge, MA: MIT Press.

Chomsky, N. 2000: *New Horizons in the Study of Language and Mind*. Cambridge: CUP.

Comrie, B. 1985: Causative verb formation and other verb-deriving morphology. In T. Shopen (ed.), *Language Typology and Syntactic Description*, Vol. III. Cambridge: CUP.

Evans, G. 1976: Semantic structure and logical form. In G. Evans, *Collected Papers*. Oxford: Clarendon Press.

Evans, G. 1980: A further reflection on semantic structure and logical form. In G. Evans, *Collected Papers*. Oxford: Clarendon Press.

Evans, G. 1981: Semantic theory and tacit knowledge. In G. Evans, *Collected Papers*. Oxford: Clarendon Press.

Gropen, J., Pinker, S., Hollander, M., and Goldberg, R. 1991: Syntax and Semantics in the acquisition of locative verbs. *Journal of Child Language*, 18, 115–151.

Higginbotham, J., and May, R. 1981: Questions, quantifiers and crossing. *The Linguistic Review*, 1, 41–80.

Higginbotham, J. 1988: Is semantics necessary? *Proceedings of the Aristotelian Society*, 88, 219–41.

Higginbotham, J. 1989: Elucidations of meaning. *Linguistics and Philosophy*, 12, 465–517.

Higginbotham, J. 1989b: Knowledge of reference, In A. George (ed.), *Reflections on Chomsky*. Oxford: Blackwell.

Higginbotham, J. 1992: Truth and understanding. *Philosophical Studies*, 65, 3–16.

Jackendoff, R. 1990: *Semantic Structures*. Cambridge, MA: MIT Press.

Keenan, E., and Stavi, J. 1986: A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9, 253–326.

Ladusaw, W. 1996: Negation and polarity items. In S. Lappin (ed.), *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell.

Leslie, A. 1995: A theory of agency. In D. Sperber, D. Premack, and A. Premack (eds), *Causal Cognition*. Oxford: Clarendon, 121–141.

Linebarger, M. 1987: Negative polarity and grammatical representation. *Linguistics and Philosophy*, 10, 325–87.

Larson, R., and Segal, G. 1995: *Knowledge of Meaning*. Cambridge, MA: MIT Press.

Ludlow, P. forthcoming: LF and natural logic. In G. Preyer (ed.), *On Logical Form*.

Montague, R. 1974: Universal grammar. In R. Thomason (ed.), *Formal Philosophy: Selected Papers of Richard Montague*. New Haven: Yale University Press.

Moravcsik, J. 1981: How do words get their meanings? *The Journal of Philosophy*, 78, 5–24.

Pesetsky, D. 1995: *Zero Syntax*. Cambridge, MA: MIT Press.

Pinker, S. 1989: *Learnability and Cognition*. Cambridge, MA: MIT Press.

Pinker, S. 1999: *Words and Rules*. New York: Basic Books.

Tanenhaus, M., Boland, J. Mauner, G., and Carlson, G. 1993: More on combinatory lexical information: thematic structure in parsing and interpretation. In G. Altmann and R. Shillcock (eds.), *Cognitive Models of Speech Processing*. Hillsdale, NJ: Lawrence Erlbaum.

Tenny, C. 1994: *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Thomson, J. 1970: Individuating actions. *The Journal of Philosophy*, 68, 774–781.

Thomson, J. 1971: The time of a killing. *The Journal of Philosophy*, 68, 115–132.