

STEPHEN G. PULMAN

*Centre for Linguistics and Philology, Oxford University,
Walton Street, Oxford OX1 2HG
Stephen.Pulman@linguistics-philology.ox.ac.uk*

LEXICAL DECOMPOSITION: FOR AND AGAINST

1. INTRODUCTION

This paper is about a problem in lexical semantics. The problem is this: to capture the validity of a variety of inferences that seem to be part of the semantic competence of a native speaker, words appear to need to be analysed as if they had internal structure. But what is the nature of this internal structure? There are at least two popular answers to this, both unsatisfactory: the first is that they are words themselves; the second that they are abstract semantic entities of some kind. The first answer is unsatisfactory because on this account we do not get all the right syntactic or semantic properties coming out. The second answer is unsatisfactory because there seem to be no empirically or philosophically satisfactory account of what these entities are.

Some influential writers, notably Jerry Fodor, have argued that the failure to find satisfactory answers to these questions shows that words do not after all have any internal structure, and maintain that in any case, the kind of inference we are trying to capture is not part of the semantic competence of the native speaker. However, while the range of meaning-constitutive inferences may be smaller than a traditional textbook on semantics might claim, I nevertheless maintain that there are some, and that some means should be found to capture them in a way that avoids objections to ‘semantic primitives’ accounts. I go on to explore a line of analysis that derives ultimately from the philosophy of action, and which sees some kinds of events as having internal structure of a kind consonant with the inference patterns we are interested in. However, although in some ways an improvement on the ‘lexical decomposition’ approach, my conclusion will be that this account is still not wholly satisfactory.

2. THE DESCRIPTIVE PROBLEM

We will begin with a brief historical review of linguistic treatments of this phenomenon. Within generative linguistics, Fillmore (1966) was perhaps the first to point out that there was a semantic regularity involving the verbs ‘come’ and ‘bring’ that was also found in a variety of other groups of verbs:

2 LEXICAL DECOMPOSITION: FOR AND AGAINST

Bill brought John.	→	John came.
Bill brought the problem up.	→	The problem came up.
Bill brought John to.	→	John came to.
He brought the truth out.	→	The truth came out.
He was brought to trial.	→	He came to trial.
? He brought the toy apart.	→	The toy came apart.
He was brought up a Mormon.	→	? He came up a Mormon.

As the latter two examples illustrate, not every verb-particle combination involving ‘bring’ and ‘come’ works this way, and there are typically some aspects of the interpretation of each member of the pair which makes us hesitate to say that the relation is one of straightforward entailment: if Bill brought John to the party unconscious in a sack, does it follow that John came to the party? or that John came to the party unconscious in a sack? I feel happier about the latter than the former, which without modification has a suggestion of voluntary agency that the circumstances do not support. This seems to be a general phenomenon with this type of inference and so from now on I shall assume that any relevant modification (temporal, locative, etc) is held constant from the implicans to the implicandum.

This ‘causative’ inference is of course widely observed in other contexts:

X killed Y	→	Y died
X melted _{transitive} Y	→	Y melted _{intransitive}
X put Y in/on/near Z	→	Y is in/on/near Z (but NB prepositions like ‘into’)
X persuaded Y to leave	→	Y intends to leave
X persuaded Y that Z will leave	→	Y believes Z will leave
X forced Y to leave	→	Y left

Although there are some exceptions, these verbs all display roughly the same pattern, and thus in order to ‘capture a generalisation’, as we used to say, it would be good to find some systematic way of deriving these inferences rather than just listing them. The earliest theory within generative linguistics, and still in many ways the dominant paradigm, is the ‘lexical decomposition’ approach, which tries to account for these inference patterns by assuming that verbs are semantically complex and can be decomposed into their component meanings. Different verbs may nevertheless share some meaning components. Note that there is an unexamined methodological linguistic assumption here that the only way to express a generalisation is by encoding it into the linguistic representations assigned to sentences.

This approach begins with the generative syntax/semanticists like McCawley 1971, Lakoff 1972, through to other ‘interpretive’ generative semanticists like Jackendoff 1990, and through to more recent and fashionable syntactic theories (Hale and Keyser, 1993, 1997). Within natural language processing and artificial intelligence

this approach to lexical semantics has also been widely adopted in one or another form: Wilks 1975, Schank 1975, and of course many of the relevant issues were discussed in Spärck Jones 1964.

X killed Y	= X caused [Y died]
X melted _{trans} Y	= X caused [Y melted _{intrans}]
X put Y Prep Z	= X caused (by ...) [Y Prep Z]
X persuaded Y to Z	= X caused (by ...) [Y intends to Z]
X persuaded Y that Z	= X caused (by ...) [Y believes that Z]
X forced Y to Z	= X caused (by ...) [X Zed]

(The (by ...) components are the place where what is specific to the action of putting, persuading etc. is meant to be detailed.) Now we can have just one inference rule capturing what these inferences have in common:

$$X \text{ cause } Y \text{ (at time } t) \rightarrow Y \text{ holds (at } t+1)$$

This rule ‘captures the generalisation’ since it applies to a natural semantic class of verbs, rather than stating similar properties over and over again for each verb.

3. FOR AND AGAINST LEXICAL DECOMPOSITION

Of course, such an analysis presupposes that we can give satisfactory answers to several questions:

Is there any independent evidence for (mono-morphemic) words having internal structure, given that *prima facie* they should be atomic? This internal structure might be covertly syntactic, or purely semantic. (I shall refer to this as the ‘internal structure’ question.)

What is the status of ‘cause’, ‘melt_{intrans}’, ‘intends’, ‘believes’, etc. in the right hand side of the analysis? Are they English words, or some abstract theoretical construct? If the latter, what is their independent justification? (The ‘ontological’ question.)

Is it possible to fill in the (by ...) component so as to give a complete analysis of the left hand side? (By ‘a complete analysis’ we mean one in which the = is interpreted as a biconditional, i.e. that the right hand side gives necessary and sufficient conditions for the truth of the left hand side, and vice versa.) (The ‘definitional’ question.)

Note in passing that this kind of lexical decomposition is rather different from that which was proposed in the days of structuralist linguistics, where words constituting lexical fields were analysed in terms of a set of n-ary features. For example, kinship terminology or folk taxonomies across languages were frequently analysed in terms of features like +/-male, +/-parent, +/-sibling etc. (e.g. Nida 1951) For this type of analysis, there is a perfectly good answer to the second question, because the features in question were generally those that could be identified in a

relatively language-neutral way, since they were perceptual, biological or social properties invariant across cultures. The first and third questions do not arise, because - at least for some of these scholars - these were not intended to be analyses of the meanings of the words involved. The intention was rather to be able to compare how different languages carved up the same 'semantic space' in partially different ways. For this of course it is essential that the dimensions of classification should be identifiable independently of any particular language, but there needs to be no claim that the words are composed of the dimensions in question.

4. THE INTERNAL STRUCTURE QUESTION

A variety of arguments have been offered suggesting that internal structure in these verbs can be discerned by using traditional syntactic tests for ambiguity. For example, it has been claimed (Lakoff, McCawley) that pronouns and ellipsis seem to be able to access the different components of a verb:

- (1) The sun melted the glass and that surprised me.

The claim is that this has two readings, one on which 'that' refers to, roughly, 'the sun melted the glass' and one on which it refers to 'the glass melted'.

- (2) The sun melted the glass and I was surprised that it did.

Similarly, here the ellipsis is supposed to be interpreted as either 'surprised that it (the sun) melted the glass', or 'surprised that it (the glass) melted'. These judgments are highly variable: I personally do not share them. Nor do they extend easily to other pairs of verbs putatively involving a similar relationship. Whereas I can just about interpret 'that' as the proposition that 'Bill died' in the first example,

- (3) a. John killed Bill and that surprised me.
b. John killed Bill and I was surprised that he did.

I find it difficult to interpret the second conjunct of the b example as meaning 'I was surprised that Bill died'.

A more convincing set of arguments for internal structure are those based on adverbial modification, it seems to me:

- (4) a. John almost killed Bill.
b. John almost did something that would have killed Bill.
c. John did something that almost killed Bill.

Example a. has two interpretations, paraphrasable as b. and c. If 'killed' has the internal structure suggested above then there is a nice explanation for this in terms of differing scopes for the adverb, suggested informally by:

X almost caused [Y died]
 X caused [Y almost died]

A somewhat similar argument is based on adverbs like ‘again’:

- (5) a. John closed the door again
 b. John closed the door before
 c. The door was closed before

The example is ambiguous in that what is asserted to be happening for the second time is either John’s action of closing the door, or the door being closed (not necessarily by John on the earlier occasion). Again we can describe this by scope:

again [X cause [Y closed]]
 X cause [again [Y closed]]

Note also that independently of the causal properties of these particular verbs, there are more general arguments for verbal predications having a complex denotation consisting of various subevents, from the literature on aspectual modification.

- (6) a. Joe was building a house
 b. Joe hired a car for a day

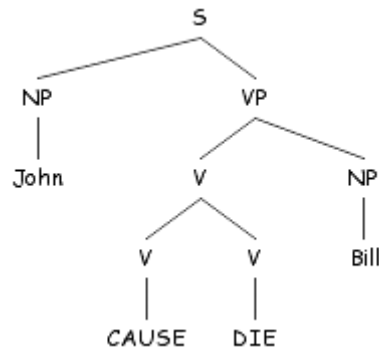
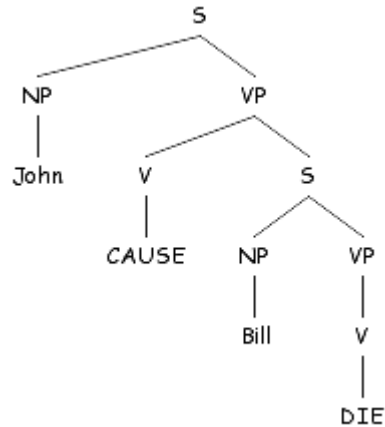
In the ‘imperfective paradox’ type sentences as in (a), the VP ‘building a house’ is interpreted as describing a process which will result in a state, but when combined with the progressive only the preliminary process is asserted to hold. In the (b) example, the temporal modification most plausibly holds of the resulting state (i.e. Joe had the car for a day), but with a strange context and an iterative interpretation could hold of the initial event or process that results in that state. Either way these observations have typically been taken to show that the events these sentences describe have internal structure. (However, some more argumentation is perhaps required to show that this is associated with the verb independently of its arguments and modifiers.)

THE ONTOLOGICAL QUESTION

The most carefully argued answer to the ontological question was perhaps that given by Lakoff and the ‘Generative Semanticists’: the view argued for by Lakoff (1972) is that the abstract components ‘cause’, ‘die’, ‘intend’ (traditionally rendered in upper case: CAUSE, INTEND, DIE etc.) are abstract arbitrarily named predicates in a ‘natural logic’. They may look just like English words in capital letters, but could instead be written ‘ABC123’ or ‘21s98j’, or any other distinct symbol. They correspond to universal concepts, frequently realised as affixes or auxiliaries in other

languages (i.e. they are readily grammaticalised: this point presumably applies more to notions like 'cause' than 'die').

Words are derived from them by processes analogous to syntactic rules:



Then a process of 'lexical insertion' replaces the complex [CAUSE DIE] by 'kill'.

A more recent adherent of this type of theory, without the syntactic claims, is Jackendoff (e.g. 1990), for whom semantics is the study of the internal mental representations that constitute 'conceptual structure' and the formal relations between this level and other levels of representation. Jackendoff assumes a set of what we might call 'conceptual parts of speech': Thing, Event, State, Action, Place, Path, Property, Amount... etc. Syntactic constituents translate into things of these types. Particular words add extra more detailed 'concepts' which are taken as primitives of our internal mental representations. Thus an apparently simple example like:

- (7) Joe drank the wine

involves an analysis of ‘drink’ as constituted of the complex:

[*Event* CAUSE([*Thing*]i,
 [*Event* GO([*Thing* LIQUID]j
 [*Path* TO([*Place* IN([*Thing* MOUTH OF ([*Thing*]i))))))]])]

‘Inference rules’ can be defined which will be triggered by such structures:

At the termination of [*Event* GO(X,[*Path* TO(Y))]
 it is the case that [*State* BE(X,[*Place* AT(Y))]

Jackendoff gives no explicit answer to the question of how GO, AT etc are to be interpreted. He says they are theoretical constructs that help to explain our intuitions about meaning. Elsewhere he claims that meaning postulate based theories like those of Lakoff are ‘notational variants’ of his, so presumably he would subscribe to the view expressed by Lakoff much earlier. (More on meaning postulates below.)

More recent theories that endorse abstract lexical components as part of the ‘abstract syntax’ of verbs, notably Hale and Keyser 1993, typically do not even address the question of the status of the ‘words’ (or perhaps ‘concepts’) they are dealing with. Fodor and Lepore (1998) point out that their invocation of more sophisticated syntactic models does not absolve them from an account of their primitives and suggests that they are subject to the same criticisms as earlier lexical decompositionists.

There are several well known problems with the Lakoff/Jackendoff/Hale-Keyser view of the status of the abstract predicates invoked by their theories. Let us take the various options one by one: firstly, of course, we could just interpret things like ‘CAUSE’, ‘LIQUID’, etc. as what they appear to be, namely English words in capital letters. Lakoff and Jackendoff do not do this, presumably because of Fodor’s (1973) arguments against this position. If these things are really just English words in capital letters, then ‘John killed Bill’ ought to be roughly synonymous with ‘John caused Bill to die’. But this is not obviously so:

- (8) a. John caused Bill to die on Sunday by poisoning him on Saturday.
 b. John killed Bill on Sunday by poisoning him on Saturday.
 c. John caused Bill to die by swallowing cyanide.
 d. John killed Bill by swallowing cyanide.

Example (a) seems easily interpretable, whereas (b) seems almost contradictory: at the very least, it needs some special accommodation to interpret, unlike (a). Example (c) is ambiguous, in that although Bill is the more likely cyanide swallower, it is linguistically possible to interpret the sentence with John swallowing

cyanide (thereby dying and releasing the handle of the trapdoor through which Bill plummeted to his death). But example (d) cannot be interpreted with Bill as the swallower of cyanide. One would pre-theoretically expect (c) and (d) to have identical linguistic properties if they are alternative realisations of the same abstract structure.

It is quite clear that there are few, if any, completely convincing cases of the synonymy that we should expect to find. Note, however (we shall return to this later) that there is no question that 'John killed Bill' entails 'John caused Bill to die': it is the reverse entailment that seems to be problematic.

The second option would be to accept Jackendoff's assertion that these predicates are theoretical constructs that (by implication) need no other justification than that they account for our intuitions. But I do not find this view acceptable, on several grounds. To begin with, Jackendoff's analysis of 'drink' above, does NOT actually account for my own intuitions about the meaning of drink, for it seems to entail that sentences like 'some reptiles drink through their pores' would be logically contradictory. But more importantly, even if your intuition is that part of the meaning of 'drink' is that liquid should enter a mouth, then unless there is some explicit connection between the construct MOUTH and the meaning of the English word 'mouth', that intuition is not accounted for. Otherwise the analysis of 'drink' above would be equally applicable to a concept which involved ingesting liquid through the ears. Unless that connection can be supplied, in a non-circular way, then the 'theoretical construct' claim is just a roundabout way of saying that the meaning of drink is whatever it has to be to satisfy your intuitions about its meaning, which does not advance things very far.

There is another way of making Jackendoff's 'theoretical construct' claim, and another way of rebutting it, which is, I find, illuminating of the issues. (See Evans and McDowell 1976). I repeat it here because it does not seem to be an argument that many linguists are familiar with - at least, their continued adherence to upper-case-ese suggests not.

On this construal, notions like 'CAUSE', 'LIQUID', 'MOUTH' etc are terms of a semantic metalanguage constituting part of the theory of meaning (for English). Any theory has to be stated in some language, preferably formal enough to see the consequences of it, and this is that language. It provides an account of the semantic intuitions of the native speaker in something like the way the theory of syntax does for intuitions of grammaticality, ambiguity, etc, since the speaker is claimed to have tacit knowledge of this theory.

Now, many people have found the notion of 'tacit knowledge' problematic. But surely one test of the adequacy of a claim that some kind of tacit knowledge of a theory underlies an ability is that explicit, non-tacit knowledge of that theory should be sufficient to exercise the ability in question. In the case of most theories of syntax, at least where they are formalisable, this condition is satisfied: if you could explicitly memorise and manipulate the various processes and representations involved, then when presented with various sentences from the language in question you can say whether it is grammatical, ambiguous, etc. But in the case of a theory like Jackendoff's, if you were able to memorise the various representations and processes, then all that you would be able to do when presented with a sentence is to

translate it into the appropriate expression of the metalanguage. Given some of the ‘inference’ rules that Jackendoff provides, you might be able to check that the representation of one sentence was derivable from that of another. But what you would not be able to do is to determine whether a sentence was true or false, or even describe the conditions under which it would be true or false. To do this you would have to know what the truth conditions of the semantic metalanguage expressions were, and the ‘theoretical construct’ claim does not provide this. (Notice also that although the inference rules look as if they enable you to check for entailments, in fact, they do not, for the semantic correlate of one representation being derivable from another is not made explicit via such rules, strictly speaking. The sample rule given is quite consistent with the two representations being synonymous, for example. To make these distinctions reference has to be made to the properly semantic properties of the representations using notions like ‘truth’, so that synonymy can be distinguished from entailment).

To summarise, then, unless we have some independent idea of what these ‘theoretical constructs’ mean, all that is achieved by a theory like Jackendoff’s is a translation from representations whose semantics we have intuitions about (namely English sentences) to representations for which we have no interpretation, not even an intuitive one. As many have remarked, this just replaces one mystery by a deeper one.

The third option, Lakoff’s, is that the meanings of these predicates are fixed by definitions or meaning postulates. Meaning postulates, as introduced by Carnap, were expressions of first order logic of the form $\forall x.P(x) \rightarrow Q(x)$, i.e. universal quantification over a material conditional. The meanings of P and Q are assumed to be known already (in the unenlightening way that predicates are interpreted in first order logic, via an assumed interpretation function that associates predicates with the appropriate sets, sets of tuples, etc.) and the function of the meaning postulate is to restrict the class of models to be considered as possible for the first order theory in question, namely those models which were consistent with the postulates. Meaning postulates do not define new concepts, they relate existing ones. So for the claim that concepts like ‘CAUSE’ and ‘INTEND’ etc are defined by meaning postulates to make sense, what must be meant is that these concepts are defined (somehow, presumably via some biconditionals) and then related by postulates.

But how are these concepts to be defined? When we give a definition of one concept in terms of another, we presuppose the defining concept to be antecedently understood. For example, if we understand addition (+) and subtraction (-), we can define multiplication (*):

$$N*1 = N \text{ (base case), and } N*M = N+(N*(M-1)) \text{ (recursive case)}$$

We can also define concepts to have certain properties even if it is difficult to attach intuitive significance to them: e.g.

$$\text{infinity} + n = \text{infinity, for any } n.$$

But how would we get started in defining ABC123 in terms of other similar predicates? At least one of these predicates would have to be understood already, otherwise we would have a completely free-floating structure of concepts linked to each other but not connected to anything external. For a predicate to be understood already, it would presumably have to correspond to a word or a phrase of a language (for that is all we have semantic intuitions about) but that would make it rather difficult to claim that the defined predicates correspond to universal, language-independent concepts.

Finally, note that Dowty 1979 does make an attempt to provide truth conditions for abstract predicates like CAUSE in terms of concepts in interval logic. This makes the relation between CAUSE and 'cause' rather more complex, but is at least the right thing to try to do, since we are now relating CAUSE to something that is independently understood.

THE DEFINITIONAL QUESTION

In a lexical decomposition analysis, the assumption must be that the components jointly constitute a definition of the meaning of the decomposed word. If the components are themselves words, then we should have full-blown synonymy.

But as many have remarked, most persistently in recent years Jerry Fodor, full synonymy is exceedingly difficult to find, even if we are careful to observe distinctions like use and mention, object- and meta-language. The general form of counterexamples follows those quoted from Fodor earlier: given a putative definition of the form A is B, we find a situation of which B plausibly holds, but of which it would be at best strange, at worst false, to say that A holds.

(Actually, those given earlier from Fodor were running together two distinct issues: synonymy, and isomorphism of presumed underlying syntactic structure with an overt counterpart syntactic structure.)

For people like Lakoff and Jackendoff, these counterexamples presumably have little force, because CAUSE is not the same as 'cause'. They can always appeal to the extra content in the (by) clause, or the difference between CAUSE and 'cause' etc to explain the apparent lack of synonymy, much as the scientist can always accommodate small tolerances between predicted and observed measurements due to the imperfections of instrumentation. One might also observe that it is a tendency for forms similar in meaning to drift apart. Given that choice of one over the other will be apparent to the hearer (as a competent speaker) that choice will tend to have an informational value in itself, perhaps for Gricean or simply for information-theoretic reasons. That extra informational value can become conventionally attached to the relevant form, leading it to become semantically distinct.

Note again that (if there is a one-to-one correspondence between CAUSE and 'cause') although these are good arguments for non-synonymy, they do not challenge the observation that there is a relation of entailment: if X is a bachelor,

then X is unmarried; if X killed Y, then X caused Y to die; if X persuaded Y that P, then X caused Y to believe that P, etc. It is the entailment in the other direction that is the doubtful one.

FODOR ON CONCEPTS

On Fodor's Representational Theory of Mind (see, for example, Fodor 1998), mental representations are syntactic objects, expressions of a 'language of thought', but with a semantics that ensures some kind of connection with things 'out there'. Mental representations of this type play a causal role in behaviour, and any putative psychological laws governing human behaviour will quantify over them. Thought is a kind of computation: syntactic processes operating on symbols, but in a way that preserves and respects the semantic properties of these symbols. The question of how these symbols get their semantic properties is thus crucial for Fodor: computation has to be more than just the shuffling of symbols. Fodor considers and rejects several different types of theory of content. The one which corresponds to one version of the lexical decomposition tradition is essentially that on which complex concepts are built up of simpler atomic ones, which are perceptually based if you are a traditional empiricist, or innate in some other way if you are a rationalist. Fodor rejects this theory on the grounds that if any version of it were true, there would be more compelling examples of adequate definitions, analyticity, full synonymy etc, than there actually are, as we have already seen.

Fodor likewise rejects what he calls 'inferential role' theories; theories that would see concepts as deriving their content from their role in a network of inferentially related concepts (perhaps via meaning postulates), something like Quine's 'web of belief'. The concepts are partially learned empirically in some way, but their full content is derived from their place in the inferential network. Again, he maintains, on this theory one would expect to see more examples of satisfactory definitions and analyticity.

Fodor has other more ideological reasons for rejecting this latter kind of theory: if the content of concepts is defined via their place in a network, and if this network derives even in part from empirical experience, then there is no guarantee that networks or concepts will be identical across individuals. Thus mental representations across people will not be type-identical, and so there will be no psychological laws quantifying over them. Since Fodor is committed to there being a level of explanation for behaviour which is irreducibly psychological or cognitive, this is not a conclusion he can accept. Furthermore, if there are no cross-person identities stable enough to support nomic generalisations, the door will be open to the worst varieties of relativism: one person's way of carving up the world will be no more privileged or 'correct' than any other person's. But while it is a reassuring fact that no two people see the world in exactly the same way, it is still the case that the similarities are vastly more striking than the differences, pathologies aside (if

relativism is true, we have no basis for calling them pathologies). Notice that neither of these objections apply to the ‘semantic primitives’ theory, at least if the primitives are perceptually based, or innate, because we can then be sure that there are some points of identity across people independently of their exposure to the empirical world.

Fodor eventually arrives at a rather radical position concerning concepts (i.e. word meanings). The lack of convincing examples of definitions, etc. leads him to the claim that all (mono-morphemic) concepts are atomic. The ideological requirement to maintain that psychological laws exist and to avoid relativism leads him to the view that, in some sense therefore, all atomic concepts must be innate, since they have no more primitive components to be constructed out of. However, the actual sense in which all concepts are innate turns out to be rather weak. Faced with the absurd possibility that concepts like ‘doorknob’ or ‘spade’ might be innate, Fodor claims instead that what is innate is the ability to form concepts of that kind on the basis of rapid and variable exposure to instances. ‘Concepts of that kind’ means concepts that are identified via properties that depend on us - as opposed to natural kind terms (water, tiger, etc.), the content of which is semantically perceived to depend on some ‘essence’ particular to that kind. The latter are innate in an analogous way: the claim (see Laurence and Margolis, 2002) is that this semantic perception of shared ‘essence’ as the basis for membership of a kind is hard-wired in us.

The claim that concepts are atomic also requires Fodor to claim that whatever inferences might be associated with concepts (linking them to other concepts), they are not part of their meaning, but something that is learned separately and which is not necessary for possession of the concept. For example, he argues, the number two is both even and prime, but surely neither of these properties need be known for the concept ‘two’ to be understood. Or, more radically, if X is a dog then X is an animal. But surely, he says, someone could understand ‘dog’ and ‘animal’ perfectly well without making this connection, and once the connection was made, we would not want to say our concepts had changed. (Consider the proposition that ‘dinosaurs were reptiles’, a proposition firmly instilled in me by my children’s encyclopedia many years ago. I now believe that dinosaurs were warm-blooded and perhaps the ancestors of birds. But I am still talking about the same creatures in these two semantically incompatible sentences.)

Here, then, is our dilemma. While I find many of Fodor’s arguments against meaning-constitutive inferential relations convincing, I also find that there are many examples where the argument is dramatically less so. It seems to me that many inferences (although perhaps fewer than we used to think before Kripke and Putnam) are just not like the ‘two’ or ‘dog’ examples. If you don’t know that ‘X killed Y’ implies ‘Y died’, or that ‘X persuaded Y to/that P’ implies ‘Y intends to/believes that P’, or that ‘X melted the chocolate’ implies ‘the chocolate melted’, then you simply won’t be able to use these words properly. You will assent to the consistency of sentences or exchanges like:

A: X killed Y, you know.

B: But Y is alive!
 A: Yes, I know that.

Whereas it may be just about conceivable to discover that our taxonomies of the natural world are so wrong that ‘dogs are animals’ has to be regarded as false (and hence cannot be partly constitutive of the meaning of ‘dog’) it just does not seem so plausible that we could empirically discover that it was after all possible to persuade someone that P without them coming to believe that P, that we could discover that someone could melt the chocolate without the chocolate melting. Notice that the ‘dogs are animals’ cases actually need a lot of work: how plausible is it that we could go on to discover that even ‘dogs are things’ was false? If there are some meaning-constitutive relations between words, then there are some analytic inferences, although there need not be any definitional equivalences. Although all persuadings of X that P involve communicating something that causes believings by X that P, it is not the case that all communications causing such believings count as persuadings. Fodor writes as if he believes that any kind of meaning-constitutive inference must imply the existence of definitional equivalences, but this is not so unless further assumptions are made. For example, any hierarchical taxonomy offers an example of a system of inferences that without further adornment supports no definitions. The inferences are all one way (transitive ‘is a’ inferences). It is only when logical vocabulary like ‘not’ and ‘or’ and a closed world assumption are added that we can go on to construct equivalences or definitions.

HORNSBY’S ANALYSIS

In the course of the development of a theory of action, Hornsby (1980: Appendix A) offers an account of inferences similar to some of those we have been discussing which avoids many of the problems of the lexical decomposition approach. In Hornsby’s analysis, the transitive and intransitive versions of verbs like ‘break’, ‘melt’ etc are related by an equivalence of the form:

$$9 \quad \exists e.V_{\text{trans}}(\text{Subj}, \text{Obj}, e) \leftrightarrow \exists e\exists f.action(\text{Subj}, e) \wedge cause(e, f) \wedge V_{\text{intrans}}(\text{Obj}, f)$$

where e and f are Davidsonian event variables. (Subj and Obj are placeholders for the corresponding variables or constants). The predicates ‘action’ and ‘cause’ are to be interpreted as the corresponding English words. (Note that this makes causation a relation between events: a position argued against very persuasively by Mellor (1987), but we will ignore that large question here).

Of course, this equivalence immediately falls foul of the Fodorean non-synonymy objection: Hornsby adds to the right hand side the further requirement:

$$10 \quad \dots \forall g\forall x.g \neq e \wedge action(x, g) \wedge cause(g, f) \rightarrow cause(g, e)$$

i.e. ... if every event other than e which causes f also causes e. The idea of this is to reduce the length of the causal chains which can count as instances of the transitive form. If John sank the boat, then the boat sank. But if John did some action A that caused the boat to sink, it will only be the case that John sank the boat if everything else that caused the boat to sink also caused A. So if John told Bill to make a hole in the boat, it will not follow that John sank the boat, even though he caused the boat to sink, because Bill making a hole did not cause John to tell him to do that. But if John hit a nail with a hammer, and via this made a hole in the boat, and this hole caused the boat to sink, then John did sink the boat, because hitting the nail caused the hole.

It is not quite clear how to interpret Hornsby's theory as a piece of linguistics. There is little hope of deriving logical forms like those above while observing any reasonable version of compositionality. It is best to regard something like 9 as a second-order schema which will produce equivalences for each of the relevant class of verbs allowing the inferences we want to go through. This strikes me as an equally good way of 'capturing a generalisation' without encoding it directly in the logical form of the sentences.

ARE EVENTS ATOMIC? PIETROSKI'S ANALYSIS.

However, as many people have noted, most recently Pietroski (1998), causative analyses like these suffer from a problem. If we take examples like:

- 11 a Booth pulled the trigger (with his finger, on April 13th)
 b Booth shot Lincoln (with a pistol, on April 13th),

it seems plausible that these sentences are both made true by the same event, the event that would be the action in Hornsby's analysis. And this event will be what the modifiers 'with NP' and 'on NP' are predicated of. But the pattern of entailments does not display the extensionality one would expect if this were so: it does not follow that:

- 12 a Booth pulled the trigger with a pistol.
 b Booth shot Lincoln with his finger.

Pietroski restricts his discussion to cases like 'X boiled/melted Y, so Y boiled/melted' (although he also includes non-homonymous pairs like 'raise' and 'rise', and even 'kill' and 'die', the latter on etymological grounds). The main claim that Pietroski makes is that we should see the events described by sentences like these as complex: while in the Booth-Lincoln cases above there is in some sense only one action, different events can be 'grounded' by the same action. Furthermore, one event may be the beginning part of another: it may be possible to discern several different 'culminations' of the same action.

The relations of grounding and culminating are defined by assuming a notion of causation, and a mereological structure on events using the notion 'part of'.

Event A ‘grounds’ event B iff A and B occur, A is a (possibly improper) part of B, and A causes every event that is a proper part of B but not a part of A (1998: 81).

Event B ‘culminates’ event A iff A and B occur, B is a (perhaps improper) part of A, and B is an effect of every event that is a proper part of A but not a part of B. (1998:86) (Every event grounds itself, and every event terminates itself, on these definitions, for (say) in the first definition, if A is an improper part of B (i.e. $A = B$), there will be no events that are proper parts of B but not parts of A for A to cause, and so the second conjunct will be vacuously true: analogously for the second definition.)

But one subevent may ground different culminating events, allowing Pietroski to give an analysis of what is happening in sentences like:

- 13 a Booth pulled the trigger (with his finger, on April 13th).
 b Booth shot Lincoln (with a pistol, on April 13th).
 c Booth killed Lincoln (on April 13th?)
 d Lincoln died (on April 14th).

On Pietroski’s analysis, the same subevent (an action of Booth) can ground three distinct events, with different culminating subevents. So the modifiers are not predicated of the same events, and the entailment puzzle concerning the a and b examples goes away. Furthermore, if ‘Booth shot Lincoln on Tuesday’ is true, but Lincoln did not die until Wednesday, it will not be true that ‘Booth killed Lincoln on Tuesday’ or that ‘Booth killed Lincoln on Wednesday’, because there will not be an event containing the right grounding and termination wholly temporally contained in either of those days. Note that this analysis requires it to be a property of events that given a cause and a related effect, there is not necessarily a single event containing both. Otherwise, if John burned the house down, and during this event, a pan of water boiled, it would be true that ‘John boiled the water’, which does not seem to reflect intuition (although ‘John caused the water to boil’ or ‘John caused something to boil the water’ would both be true).

Returning to our causative inferences, the neo-Davidsonian logical forms that Pietroski assigns to the $melt_{trans}/melt_{intrans}$ sentences are:

- 14 a Nora melted the chocolate.
 b $\exists e.melting_{trans}(e) \wedge Patient(e,the\ chocolate) \wedge Agent(e,Nora)$
 c The chocolate melted.
 d $\exists f.melting_{intrans}(f) \wedge Patient(f,the\ chocolate)$

The notions of Agent and Patient (partly) are defined in terms of grounding and culmination:

- 15 a $\forall e. \forall x.Agent(e,x) \leftrightarrow \exists a.grounds(a,e) \wedge action(a,x)$
 b $\forall e. \forall x.Patient(e,x) \leftrightarrow \exists f.culminates(f,e) \wedge Patient(f,x)$

Substituting in these definitions, the logical form for ‘Nora melted the chocolate’ will now be equivalent to:

$$16 \quad \exists e.\text{melting}_{trans}(e) \ \& \ \exists f.\text{culminates}(f,e) \ \& \ \text{Patient}(f,\text{the chocolate}) \\ \& \ \exists a.\text{grounds}(a,e) \ \& \ \text{action}(a, \text{Nora})$$

Pietroski takes the intransitive form of the verb to be basic and defines the transitive form:

$$17 \quad \forall e.\text{melting}_{trans}(e) \leftrightarrow \exists f.\text{melting}_{intrans}(f) \ \& \ \text{culminates}(f,e)$$

Given this definition the logical form of ‘the chocolate melted’ is now entailed by that of ‘Nora melted the chocolate’.

EVENT ONTOLOGY

Pietroski’s analysis gives us an account of the entailment relations between sentences involving transitive and intransitive forms of a verb, without committing us to abstract causative morphemes. Could we extend this approach to our ‘persuade/intend’ or ‘kill/die’ examples? To try to do this we need to get a little clearer about what model theory is presupposed by the logical forms above.

What exactly now are the denotations of verbs like ‘break_{trans}’ and ‘break_{intrans}’? On Pietroski’s analysis, any event which culminates in a ‘break_{intrans}’ event will count as a ‘break_{trans}’ event. But every event culminates itself, and so every ‘break_{intrans}’ will count as a ‘break_{trans}’ (1998:105). Pietroski makes use of this fact to deal with cases where the ‘break_{trans}’ does not have an agent in the usual sense, but for now the important point is that the transitive denotations (properly) include the intransitive denotations.

To put a little more detail on this we will adopt the model theory proposed for events in Kamp and Reyle (1992) DRT, where the denotation of an n-place verb is given in terms of an n+1 tuple of individuals in the domain, of which the first (conventionally) is an event. So the denotation of ‘snore’ is a set of <Event,Individual> pairs such that the first member of the pair is an event of snoring by the second member, and so on for transitive and ditransitive verbs.

Under a version of this proposal, a sentence like ‘The water boiled’ is made true by the existence of an intransitive boiling event of which the patient is ‘the water’. A sentence like ‘John boiled the water’ is made true by the existence of a complex event including two subevents, the grounding event which is some action by John (which does not necessarily correspond to a linguistic denotation), and the culminating event, which is a boiling_{intrans} event. Let the notation ‘event_{grnd/culm}’ mean an event which is a grounding or a culmination event, and let the notation

‘event_{<grnd...culm>}’ mean an event which contains both a grounding and a culmination event. ‘Ind’ means a non-event individual which will be subscribed by agent or patient. So the denotation D of intransitive and transitive ‘boil’ etc. will be:

D(boil_{intrans}) = a set of <event_{culm,ind_{pat}}> pairs
 D(boil_{trans}) = a set of <event_{<grnd...culm>,ind_{ag},ind_{pat}}> triples, with the constraint that the extractable tuples <event_{culm,ind_{pat}}> are in D(boil_{intrans})

It is straightforward to see how this would extend to the ‘kill/die’ cases:

D(die_{intrans}) = a set of <event_{culm,ind_{pat}}> pairs
 D(kill_{trans}) = a set of <event_{<grnd...culm>,ind_{ag},ind_{pat}}> triples, with the constraint that the extractable tuples <event_{culm,ind_{pat}}> are in D(die_{intrans})

Notice that in these definitions, the culminating events are constrained to be in the denotation of the intransitive verbs. The grounding events, by contrast, are not. There is no requirement in these verbs for the set of grounding events characteristic of killing or melting to correspond to the denotation of any other word or phrase. But perhaps for other classes of verb, there is: perhaps we could extend this analysis to our ‘persuade’ cases by requiring the grounding events to be of a particular kind. Verbs like ‘persuade’, for example, seem to require that whatever it is brings about the beliefs or intentions of the persuadee, it has to be some event of communication. Analogously, the culminating events need to be included in the denotation of the relevant propositional attitude. So on this kind of analysis, ‘persuade’ would denote a complex event containing at least two subevents, one, the grounding action by the agent being constrained to be in the denotation of ‘verbs of communication’, and the culminating event being constrained to be in the denotation of ‘believe’ or ‘intend’ events, depending on whether there is a finite or an infinitival complement.

D(persuade_{trans}) = a set of <event_{<grnd...culm>,ind_{ag},ind_{pat},proposition}> tuples, with the constraint that tuples <event_{grnd,ind_{ag}}> are in D(‘verbs of communication’) and <event_{culm,ind_{pat},proposition}> are in D(believe) or D(intend)

Now the entailments that we want will follow directly from the denotation of the verbs: we do not need to posit abstract ‘believe’ or ‘intend’ morphemes in the lexical representation or logical form for ‘persuade’: the relevant information is there in the denotation. In this particular case we could perhaps go on to tell a plausible developmental story about the acquisition of concepts like ‘persuade’. We would predict that until you have learned which events count as acts of communication, and which count as beliefs or intentions, you will not be able to learn which events

are in the denotation of ‘persuade’ since this includes both of the other types of event.

However, while moving responsibility for the validity of inferences from the representational level to the denotational level, we have ironically enough now got a story about the relevant entailments that is rather parallel to that we would tell for the ‘dogs are animals’ cases. Consider how our model theory would capture the ‘dogs are animals’ intuition. It would do so by stipulation that in every valid model for English, the set of dogs is to be a subset of the set of animals. Model theory for our analysis of causative verbs is of a similar nature, but using ‘part of’ rather than ‘member of’: ‘X melted Y’ entails ‘Y melted’ because whenever there is an event making ‘X melted Y’ true, there is a culminating sub-event making ‘Y melted’ true. ‘X persuaded Y to leave’ entails ‘Y intends to leave’ because the denotation of ‘persuade’ includes complex events whose culminations are ‘intendings’ by the patient. Unless we were to argue that mereological relations like ‘part of’ somehow are intuitively ‘more necessary’ than those like membership or set inclusion, we no longer have a formal basis for our intuition that there is a difference between the ‘dogs are animals’ cases, and the ‘persuade/intend’ ones. So that as well as acknowledging that with some effort you could imagine it turning out that we had misplaced dogs in our system of classification, we are also apparently acknowledging that we could discover that, after all, you could persuade someone to do something without them thereby intending to do it, since learning the meaning of ‘persuade’ is now just a question of learning its place in our ontology of different kinds of composite events. But this is not a conclusion that I am happy with.

Acknowledgements: earlier versions of this paper have been presented to audiences at the University of Edinburgh, and the University of Oxford, and I am grateful the these audiences for their comments. I have also benefitted from comments on a draft by Ted Briscoe, Nicholas Shea, and Yorick Wilks.

References

- D. Dowty 1979 *Word Meaning and Montague Grammar*, D. Reidel, Dordrecht.
- G. Evans and J. McDowell, 1976, editors’ introduction to “Truth and Meaning: Essays in Semantics” ed. Gareth Evans, John McDowell, OUP 1976
- C. J. Fillmore, 1966, “Deictic categories in the semantics of ‘come’”, in *Foundations of Language*, 2, 219-227.
- J. A. Fodor, 1970, “Three reasons for not deriving ‘kill’ from ‘cause to die’”, *Linguistic Inquiry* 1:4, 429-438
- J. A. Fodor, 1998, “Concepts”, Oxford University Press.
- J. A. Fodor and E. Lepore, 1998, “The Emptiness of the Lexicon”, *Linguistic Inquiry* 29:2 269-288.
- K. Hale and S. J. Keyser, 1993, “On argument structure and the lexical representation of semantic relations”, in “The view from Building 20”, ed. S. J. Keyser and K. Hale. Cambridge, MA, MIT Press.

- K. Hale and S. J. Keyser, 1997, "On the complex nature of simple predicators." in 'Complex predicates', ed. A. Alsina, J. Bresnan and P. Sells. Stanford, CSLI Publications: 29-65.
- J. Hornsby, 1980, "Actions", London: Routledge.
- R. Jackendoff, 1990, "Semantic Structures", MIT Press
- H. Kamp and U. Reyle, 1992, "From Discourse to Logic", Kluwer Academic Publishers.
- G. Lakoff, 1972, "Linguistics and Natural Logic", in Semantics of Natural Language ed. D. Davidson and G. Harman, D. Reidel Publishing
- S. Lawrence and E. Margolis, 2002, "Radical Concept Nativism, Cognition 86, 25-55.
- E. Lepore and J. Fodor, 1998, "Morphemes Matter," (with J. Fodor), forthcoming; RuCCS Tech Report .
- J. D. McCawley, 1971, "Prelexical Syntax", reprinted in "Semantic Syntax", ed.
- P. Seuren, 1974, Oxford Readings in Philosophy, OUP
- D. H. Mellor, 1987, "The singularly affecting facts of causation", in Metaphysics and Morality: Essays in honour of J. J. C. Smart, ed. P. Pettit et al., Oxford, Blackwell: 111-136.
- E. A. Nida, 1951, "A system for the description of semantic elements", Word 7, 1-14.
- P. Pietroski, 1998, "Actions, Adjuncts and Agency", Mind, 107 pp 73-112