PHILOSOPHICAL PSYCHOLOGY, VOL. 12, NO. 2, 1999

# In defense of definitions

DAVID PITT

ABSTRACT *The arguments of Fodor, Garret, Walker and Parkes [(1980) Against definitions,* Cognition, 8, *263–367] are the source of widespread skepticism in cognitive science about lexical semantic structure. Whereas the thesis that lexical items, and the concepts they express, have decompositional structure (*i.e. *have significant constituents) was at one time "one of those ideas that hardly anybody [in the cognitive sciences] ever considers giving up" (p. 264), most researchers now believe that "[a]ll the evidence suggests that the classical [(decompositional)] view is wrong as a general theory of concepts" [Smith, Medin & Rips (1984) A psychological approach to concepts: comments on Rey,* Cognition, 17, *272], and cite Fodor* et al. *(1980) as "sounding the death knell for decompositional theories" [MacNamara & Miller (1989) Attributes of theories of meaning,* Psychological Bulletin, 106, *360]. I argue that the prevailing skepticism is unmotivated by the arguments in Fodor* et al. *Fodor* et al. *misrepresent the form, function and scope of the decompositional hypothesis, and the procedures they employ to test for the psychological reality of definitions are flawed. I argue, further, that decompositional explanations of the phenomena they consider are preferable to their primitivist alternatives, and, hence, that there is prima facie reason to accept them as evidence for the existence of decompositional structure. Cognitive scientists would, therefore, do well to revert to their former commitment to the decompositional hypothesis.*

## 1. Introduction

The thesis that some words are semantically structured has been put to significant use in philosophy, logic and linguistics [1]. In philosophy, it has been taken to underwrite the traditional distinction between analytic and synthetic judgments, and, thereby, claims to one kind of non-empirical knowledge. In logic, it has been used to provide an account of the validity of certain non-logical inferences, and the necessity of certain non-logical truths. In linguistics, it has featured in explanations of certain semantic properties and relations of natural language expressions, as well as facts concerning language processing and understanding.

In spite of its evident utility, however, the thesis has been abandoned by most contemporary philosophers, logicians and linguists. In philosophy and logic, Quine (1966, 1980a, b) famously argued that the likeliest candidates for making objective, non-circular sense of definitions fail. A great deal of philosophy has been done on the conviction that his conclusion generalizes. In cognitive science, the arguments of Fodor, Garrett, Walker and Parkes (henceforth "FGWP") (1980) have come to be

*David Pitt, Department of Philosophy, University of Nebraska–Lincoln, USA.*

0951-5089/99/020139–18 © 1999 Taylor & Francis Ltd

seen as the definitive disconfirmation of the once prevalent decompositional approach to language structure [2]. It is now widely assumed by researchers in these areas that lexical items are, uniformly, semantically primitive.

It is safe to say that for most philosophers, logicians and linguists, the case against decompositionality rests on one of these two arguments [3]. They are the basis for the widespread rejection of lexical semantic structure. Yet there is reason for serious doubt that these arguments are as effective as they are usually taken to be. In philosophy, Katz (1967, 1988b, 1990, 1992) has argued that considerations from contemporary theoretical linguistics diminish considerably the force and scope of Quine's original arguments, and Quine has conceded the point (see Quine, 1967, 1990; Katz, 1990, pp. 199–202; Clark, 1993, pp. 7–12) [4]. It is my purpose in this paper to show that the arguments in FGWP (1980) are also insufficient to ground skepticism about decompositionality. I will argue that FGWP misrepresent the form, function and scope of decompositional theories, and that the procedures they employ to test for the psychological reality of definitions are flawed.

If these central antidecompositional arguments fail, then there is no obvious reason not to accept decompositionality. The main thesis of this paper is thus a negative one: prevailing skepticism about lexical semantic structure is unmotivated. I will, however, also try to establish a positive claim in the fourth section of the paper, namely, that decompositional explanations of the phenomena FGWP consider are preferable to their primitivist alternatives, and, hence, that there is prima-facie reason to accept them as evidence for the existence of decompositional structure.

## 2. "Against definitions" [5]

### *2.1. Reference fixing*

According to FGWP, classical theories of reference propose to explain the reference of some expressions in terms of their definitions. To give the definition of an expression, on this view, is to give another expression with the same meaning that makes the application conditions of the defined expression explicit; thus, "the definition of 'bachelor' as 'unmarried man' fixes the extension of 'bachelor' relative to the extension of 'unmarried man'" [p. 265]. The reference of the defining expression (if it is not composed of primitives) is determined by the reference of *its* definition ("adult human male without a spouse", as it might be, defines "unmarried man", which is why "unmarried man" refers to adult human males without spouses). But the process of reference determination through definition cannot continue indefinitely: eventually some expressions must be reached whose reference is not fixed by the reference of some other expressions that define them. Such expressions form the *primitive base* of the lexicon of the language. Thus, a theory of reference that makes use of definitions in fixing the reference of some terms incurs the responsibility of accounting for the reference of the primitive vocabulary. Moreover, the interpretation of primitives must proceed in a way fundamentally

different from the interpretation of the rest of the lexicon, and forms the basis for the interpretation of the language as a whole.

FGWP consider the empiricist approach to the interpretation of primitives, according to which they express sensory/motor properties. Finding such an account wanting ("there are literally *no* convincing examples of definitions which take prima facie non-sensory/motor terms into a sensory/motor vocabulary" [p. 268]), they conclude that "[i]t may well be that definition plays *no* role in theories of language and the world" [p. 284].

### *2.2. Informal inference*

A formally valid inference is one whose validity depends only on the logical form of its sentences (determined by the distribution of logical terms); an informally valid inference is one whose validity is not due to the logical form of its sentences. Another supposed virtue of definitional theories, according to FGWP, is that they reconstruct informal validity as formal validity at the semantic level. That is, the validity of an informal inference turns on the logical form of the semantic representations of the sentences involved. The validity of (1)–(2),

(1) Max is a bachelor
(2) Max is a man

for example, is due to the formal validity of (1′)–(2),

(1′) Max is a man and Max is unmarried

where (1′) is the representation of (1) at the semantic level.

The objection the authors offer here is that there are informally valid inferences, such as (3)–(4) and (5)–(6),

(3) John killed Mary
(4) Mary died
(5) The sky is blue
(6) The sky is colored

which are not formally valid at the semantic level. In the case of (3)–(4), the reason is that the definition of "kill" as "cause to die" does not license the inference—nor, it is claimed, would the definition of "cause", if there were such a thing (which is doubtful). The problem with (5)–(6) is that it is not a definitional argument; if it were, there would be a predicate "*F*" which together with "colored" would complete the definition of "blue", such that "the sky is colored and *F*" would imply and be implied by "the sky is blue". But there is no such predicate. Thus it is not generally true that definitions explain the validity of informal inferences.

FGWP propose that all informally invalid inferences be handled in logical theory by meaning postulates. Meaning postulates formalize necessary connections among predicates; thus, inferences supposedly explained by the postulation of semantic structure are handled by rules that are part of "an *enriched* inferential

apparatus" [p. 272]. There is no need for definitions in the theory of informal inference, and no need for a distinct level of semantic representation in grammar [6].

### *2.3. Sentence comprehension*

If a word has definitional structure, FGWP argue, it is natural to suppose that understanding that word involves recovering a mental representation of that structure. Mental representations are the objects over which mental operations (such as understanding) are defined. Given this, there ought to be some empirically detectable differences in the behavior of subjects in sentence comprehension tasks when there are supposed to be differences in the complexity of the definitions of constituent terms.

The authors include extensive experimental results that they take to disconfirm the definitional theory. The expected behavioral differences are not evident: definitional structure is not *psychologically real*. Since, it is supposed, grammars are theories of speakers' knowledge of their languages, definition has no place in them [7].

## 3. In defense of definitions

### *3.1. Reference fixing*

*3.1.1. Assignment and constraint.* There are two senses in which the reference of an expression may be said to be "fixed", only one of which is countenanced in FGWP's discussion. In one sense—let's call it reference *assignment*—a connection between an expression and what it refers to is somehow *established*. In the other sense—call it reference *constraint*—the referential possibilities of a particular expression are in some way *limited*.

To say that the reference of an expression is fixed relative to the extension of its definition is thus to say either that a definition *assigns* an extension to its definiens, or that it *constrains* the extension of its definiens. These functions are different: assignment is an *extra*linguistic relation, whereby language is connected to the world; constraint is an *intra*linguistic one, whereby language is connected to language in a way that limits potential connections to the world [8].

Now, the argument in "Against definitions" recognizes only the extralinguistic function of definitions. FGWP claim that "[d]efinitions [only] fix the extensions of definable expressions relative to an interpretation of the primitive basis" [p. 266], so that "[o]ne hasn't *got* a [definitional] theory of language and the world unless [the problem of interpreting the primitive base] has been adequately addressed: all one has is a theory of a relation between uninterpreted linguistic forms" [p. 267] [9]. But these claims apply only on the construal of reference fixing as reference *assignment*. Plainly, one hasn't got definitional reference assignment if the primitive basis is not referentially interpreted. But definitions *constrain* the reference of definable expressions *independently* of the interpretation of the primitive basis; so it is not the case that reference isn't constrained until the primitive base is interpreted. In the

constraining sense, "a theory of a relation between [referentially] uninterpreted linguistic forms" *is* theory enough for definitions. The failure of a theory of assignment has no bearing on the status of a theory of constraint.

*3.1.2. Empiricism.* Three points. First, to conclude that decompositionality plays no role in the theory of reference solely on the basis of the failure of the *empiricist* account of the reference of undefined terms seems too quick. For one thing, the empiricist account is not the only available one. The reference of primitives on the empiricist story is fixed (in the *assigning* sense) causally: roughly, primitive terms express the concepts (mental representations) tokened by sensory experience, and refer to the things thus causally connected to those concepts. One might, alternatively, take the operative relation between primitive (hence *any*) terms and their extensions to be *instantiation*: primitive terms refer to the things that instantiate the properties expressed by their associated concepts (concept tokenings need not be causally related to their extensions). Moreover, the failure of empiricism here should not be surprising. The empiricists claimed that the primitive/defined distinction coincides with the sensory/non-sensory distinction, not on the basis of any actual analysis of complex concepts, but on the basis of an epistemological position. The distinction between defined and primitive terms is itself antecedently clear: defined terms have definitions, primitive terms do not. But it does not follow from this that primitive terms should *also* have the property of expressing sensory/motor concepts. This is a substantive, *a priori* claim; why should it be true? Finally, given that the reference constraining conception of the explanatory role of definitions is the correct one, the failure of some theory of the interpretation of primitives is irrelevant to the status of decompositionality in the theory of reference.

Second, in their section on informal inference, FGWP advocate what I will call "lexical primitivism"—the view that we ought to "consider the entire vocabulary to be [semantically] primitive" [p. 276]. But notice that the objections they raise here for decompositional theories would count against such a theory as well, but on an even larger scale. If the interpretation of primitives is a problem in a definitional context (in which the number of primitives is assumed to be significantly smaller than the number of words), it will be that much more a problem for the theory that treats *all* terms as semantically primitive. In order to avoid such problems, FGWP would need a different theory of the interpretation of primitives. But if there is some other workable account of the interpretation of primitives, then the failure of the empiricist story cannot count as a reason for abandoning decompositional theories. One could use the new theory of the primitive base in a definitional context.

Third, if the important issue is *how* primitive terms get related to the world—and not *what* in the world they are related to, then the relevant aspect of the empiricist account is its claim that "the extensions of [primitives] are ... fixed by a causal account of the sensory/motor transducers" [p. 267]. There are three questions that must be kept separate here: (a) what are the primitive terms of the language? (b) what are the referents of those terms? and (c) what is the relation that holds between the primitive terms and their referents in virtue of which the former

refer to the latter? Only the third question is really relevant to a discussion of how language gets related to the world (reference fixing in the *assignment* sense). The empiricist story gives an answer in causal terms, but it is not the case that *any* theory that gives a causal account of reference fixing is empiricist in the sense at issue here [10]. What *ought to be* the focus of FGWP's criticism, then, is the causal account of reference, since this is the aspect of the empiricist theory that is relevant to its evaluation as a theory of reference in this context. The objection FGWP bring against the empiricist story, however, focuses on the analysis of primitives in sensory/motor terms (in effect, its answer to the second question). The failure of the empiricist account on *that* score is irrelevant to its status as a theory of how language gets connected to the world. Empiricism fails because sensory/motor terms are not the primitives—not because it fails to give an account of how the proposed primitive terms get their reference, or because the account it gives is wrong.

### *3.2. Informal inference*

The criticisms FGWP make of definitional reconstructions of informally valid inferences rest on three assumptions. The first is that all semantically structured words are definable [11]. The second is that the formalism of standard first-order logic is appropriate for the representation of lexical semantic structure, and, hence, that "definitional" theories of inference "propose to save the inferential apparatus of standard logic by constructing a level of linguistic representation at which definitional structure is displayed" [p. 278]. The third is that definitional theories are responsible for accounting for all informally valid inferences. It follows from this assumption that if there are informally valid inferences that cannot be explained as formally valid at the semantic level, definitional theories fail to explain what they are supposed to explain. In which case, FGWP argue, one has grounds for abandoning them.

All three of these assumptions are false. Theories that postulate "that the morphemes of a natural language typically have internal structure at the 'semantic level'" [p. 264] are not, *ipso facto*, committed to the view that all such morphemes are definable, nor need they construe the definitions of those that are as logical constructions out of predicates. It follows from the denial of these assumptions that inferences like (5)–(6) are not counterexamples to the hypothesis that decompositional structure can be appealed to to explain some inferences, and that the existence of both semantically valid arguments that cannot be reconstructed as logically valid via definitions and informally valid arguments that cannot be reconstructed as semantic is a "don't–care".

*3.2.1. Definition vs. decomposition.* One of the arguments FGWP take to show that not all informally valid arguments are definitional—namely, (5)–(6)—could equally well be taken to show that not all semantic structure is definitional. In fact, many philosophers suppose color terms to exemplify a kind of semantic structure that cannot be represented as necessary and sufficient conditions in a logical formalism.

Johnson (1921) introduced the terms "determinate" and "determinable" to characterize the relation between color terms and "color". Determinates analytically imply their determinables, but there is no "(secondary) adjective which analysis would reveal as characterizing these different (primary) adjectives" (determinates) in virtue of which "the several [determinates] are put into the same group and given the same name" (i.e. of the determinable) (Johnson, 1921, p. 176). That is, determinates are not definable, though they have semantic structure. Searle (1959, 1967) contrasts the determinate–determinable relation with the species–genus relation. A determinate entails its determinable, but "is not a conjunction of its determinable and some other property independent of the determinable" (1959, p. 143), whereas a species is definable as the conjunction of its genus and differentia [12].

Though he does not use the terminology of determinates and determinables, Katz (1972) also recognizes the distinct nature of the relation of color words to "color". He proposes representing the semantic structure of color words using what he calls "distinguishers", which are elements in his decompositional formalism that mark semantic distinctions but do not themselves have conceptual content. On Katz's view, color concepts are complex, and distinct from one another, though they are not definable [13].

In order to distinguish such terms from those that do have definitions, let us use the term "primitive" for lexical items that do not have definitions (statable necessary and sufficient conditions), and "simple" for those that have no semantic structure. Thus, though all simple terms are primitive, not all primitive terms are simple; and though all definitional theories are decompositional, not all decompositional theories need be definitional. One may agree that not *all* semantic inferences can be explained by definitions without conceding either that *none* can, or that lexical semantic structure is non-explanatory in general [14].

*3.2.2. Non-logical form.*

FGWP take definitions to be *conjunctions* of predicates:

> [q]uite generally, if an informally valid argument turns on a definition, then there will be some clause [C] that we can conjoin to the [predicate of the conclusion] which will make the corresponding bi-conditional [premise predicate $\longleftrightarrow$ 0,2 $\rightarrow$ conclusion predicate & C] true. Any informally valid argument which does not meet this condition can't be a definitional [read "semantic"] argument. [p. 272, emphasis added]

But not all decompositional theories construe complex semantic representations—whether definitions or not—as constructions out of predicates and standard logical operators; nor are they concerned to preserve the inferential apparatus of standard first-order logic.

Katz's decompositional sense theory (e.g. Katz, 1972, 1988a), Jackendoff's theory of lexical conceptual structures (e.g. Jackendoff, 1990), and the thematic role theories of generative grammarians (e.g. early Jackendoff, 1972; Higginbotham, 1985; Rappaport, *et al.*, 1993)—work rooted in the theories of Gruber (1965) and

Fillmore (1968)—for example, use non-logical formalisms to represent decompositional content; Katz and Jackendoff also formulate non-logical inference rules in terms of their respective formalisms.

Moreover, whatever the merits of these particular theories, there are reasons for thinking that a decompositional theory *shouldn't* accept FGWP's assumption that all lexical semantic structure is standard logical (much less *conjunctive*).

There are structural features of complex word meanings that are not straightforwardly capturable in standard logical notation—for example, the property of conceptual *superordination*. Consider the term "polygon", which I will take to be definable as "closed rectilinear plane figure". Intuitively, a polygon is a *kind of* plane figure, but not a kind of rectilinear thing, since being rectilinear presupposes being planar (only plane figures can be rectilinear). Thus, the concept PLANE FIGURE is contained in the concept POLYGON in a different way than the concept RECTILINEAR—namely, as a superordinate. But a standard logical formalism would represent the definition of "polygon" as either "closed plane figure and rectilinear" or (equivalently) "rectilinear and closed plane figure". Neither captures the superordinacy of PLANE FIGURE. Nor can it be adequately represented by meaning postulates, since the asymmetric entailment relation exhibited by superordinates and subordinates constitutes only a *necessary* condition on super-ordinacy [15].

The general point here is that a decompositional theory should probably avoid the approach to the explanation of informally valid inference recommended by FGWP. Appeals to decompositional structure should not be supposed to be appeals to semantically buried first-order logical structure; inferences such as (1)–(2) should not be construed as *logically* valid at a deeper level of analysis.

*3.2.3. Non-semantic inference*. If there are informally valid arguments not explicable on the basis of the definition of terms, this may be because either (1) the semantic structure involved is not definitional (as suggested above for color and, possibly, natural kind terms), or (2) the inference is not semantic (as, for example, arithmetic inferences (e.g. $> 4 \rightarrow > 3$) and geometric inferences (e.g. quadrilateral $\rightarrow$ quadrangular). In neither case would it follow that decompositional theories fail at explanation—in the first case because there is a decompositional (though non-definitional) explanation, and in the second case because decompositional theories *per se* are not responsible for the explanation of non-semantic inferences. An inference should only be counted as semantic if it may be accounted for on the basis of the semantic properties of constituent terms.

Moreover, if there are non-logical, non-decompositional sources of validity (namely, mathematical or, perhaps, "pure metaphysical facts") it is only on FGWP's third assumption that the failure of decomposition to explain an informal inference supports the view that the terms taken to license the inference ought to be treated as simple. That an inference is not semantic does not imply that the terms that putatively license it have no semantic structure, but only that the inference does not depend on that structure. Insofar as there are other facts that explain the inference, the semantic structure of the terms does not come into play.

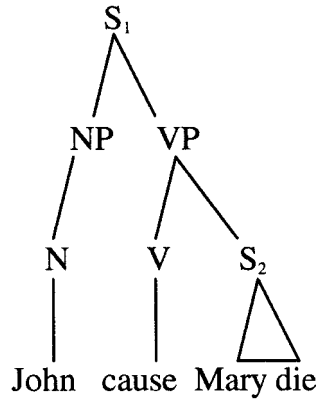



FIG. 1.

*3.3. Sentence comprehension [16]*

FGWP's assumption that grammatical structure in general must be psychologically real, and, hence, the evidential role of psycholinguistic experiments, may themselves be challenged [17]. I will not pursue this line here, however. Rather, for the sake of argument I will accept FGWP's assumption that what is not psychologically real is not linguistically real, and evaluate their procedures and arguments on their merits.

FGWP propose to test for the psychological reality of definitions in the following way. Using the example of the purported definition of "kill" as "cause to die", they give the phrase-marker in Figure 1 as the semantic representation of (3), and note that the two-clause structure induced by this analysis results in a shifting of grammatical relations between surface subject and object: "John" and "Mary" are subject and object of "kill" in the surface representation of (3), but "Mary" is the subject of "die", and no longer grammatically related to "John", at the semantic level. The strategy is to find a test procedure that is sensitive to such surface/deep shifts, validate it on non-decompositional semantic–shift sentences, and apply it to the causatives [18]. A test sensitive to such shifts that failed to detect them for sentences with causative verbs would disconfirm the proposed semantic analysis. Since causative verbs are supposed to be among the likeliest candidates for decomposition, evidence against them is good evidence against decomposition generally.

FGWP consider a number of examples of minimal pairs which they take to illustrate semantic, but non-definitional shifting. They recognize that some of the examples they use—e.g. "expect"/"persuade" and "easy"/"eager" pairs—are arguably syntactic, so they settle on (7) and (8) as representative:

(7) a. All of the men left
    b. None of the men left

(8) a. John wanted an apple
    b. John ate an apple

They claim that, in virtue of the meanings of "all" and "none", the property of having left is attributed to the men in (7a) but not in (7b); and, in virtue of the meanings of "want" and "eat", (8b) asserts a relation between John and some apple, while (8a) (on the *de dicto* reading) does not. In both cases the differences are supposedly not induced by the semantic decomposition of any word.

The test which is shown to be sensitive to the differences of (7) and (8) involves eliciting speakers' intuitions on the degree of grammatical relatedness of arbitrary word pairs in sentential contexts. Levelt (1970) concluded that such intuitions may be interpreted as providing evidence for the psychological reality of the kind of linguistic structure represented in derived phrase markers [19]. Levelt also held that such intuitions are sensitive to *underlying* sentential structure as well. For example, subjects reported the same degree of intuitive relatedness between surface subject and verb ("John"/"drove") and ellipsed (deep) subject and surface verb ("John"/"walked"), in (9):

(9) John drove to the store and walked home

FGWP conclude that, since a Levelt-style intuition test is sensitive to relations that are represented at grammatical levels deeper than surface, one ought to be able to use such intuitions to test for the predicted difference of relatedness in the pair "John"/"Mary" as it appears in "John killed Mary" and the non-shifting "John hit Mary". If the definitional analysis of "kill" as "cause to die" is correct, speakers ought to give the pair "John"/"Mary" a lower rating in the context "John killed Mary" than in the context "John hit Mary".

FGWP report that the Levelt test is sensitive to the differences in (7)–(8), and, furthermore, that it shows *no* detectable asymmetry between sentences containing causatives and sentences containing non-shifting verbs. Thus, they conclude that the definitional hypothesis is disconfirmed.

*3.3.1. A dilemma.*

But notice that the fact that the test procedure is validated on sentences whose differences are *not* definitionally induced does nothing to allay the concern that FGWP raise that a negative result might be due to the test's insensitivity to definitional shift. Their worry was that if a test procedure were tried out directly on sentences containing causative verbs, it would be unclear what the significance of a negative result would be. A positive result would, presumably, confirm the hypothesis about "abstract" structure at issue, but, since the shifts in (7)–(8) are, it is argued, non-definitional, it does not follow that a negative result would be relevant to the question of the psychological reality of definitions. One may still worry that a negative result only shows the test instrument to be insensitive to decompositional shifting, since it has only been validated on sentences that exhibit *non*-decompositional shifting.

FGWP seem to be in a methodological dilemma here. No procedure may be tested on causative sentences like (3) directly, due to the worries about negative results; but a procedure validated on *non*-shifting sentences is not thereby shown to be sensitive to the kind of semantic structure one would want to use it to test for. Either way, no significance can be assigned to a negative result.

*3.3.2. Deep structure.*

It is not clear that sensitivity to the relatedness of "John" and "walked" in (9) really is sensitivity to deep structure, or, even if it is, that this is relevant to the issues being addressed.

FGWP maintain that "[w]hat determines such intuitions is either that 'John' is the deep subject of 'walk' [20] or that 'John' is the agent of 'walk' (or both, assuming that these facts are indeed distinct)" [p. 304]. As to "John" being the deep subject of "walk", this supposes that (9) is derived from "John drove to the store and John walked home" by a deletion transformation. But the simplest hypothesis in cases like this is that "John" is the surface subject of a compound verb phrase, in which case the lack of difference between "John"/"drove" and "John"/"walked" would be explained by "John" being the surface subject of both verbs. Further, even on an analysis where (9) has an elided second "John" as subject of "drove", it is possible (and, perhaps, preferable) to construe the elision as taking place at the *phonological* level [21].

Moreover, deep structure of the kind Levelt claims his test is sensitive to is syntactic, so even if "John" were the deep subject of "walk" in the second clause of (9), Levelt's results would be relevant to the issues FGWP are addressing only on the assumption that semantic structures are deep syntactic—an assumption which has not been defended.

As to "John" being the *agent* of "walk", this relation is either syntactic (as are theta roles generally in contemporary generative syntax) or semantic (as are thematic relations generally in, for example, Jackendoff's conceptual semantics). If it is syntactic, then one cannot help oneself to the assumption that the level of grammar at which it is represented is semantic [22]. But if it is semantic, one must establish that the level at which it is represented is syntactic, or the Levelt test won't be relevant.

## 4. For decomposition

I have argued that FGWP's psycholinguistic experiments are inconclusive, at best, and that their arguments from reference fixing and informal inference rest on false assumptions. If my arguments are sound, then given the acknowledged limitation on the scope of Quine's arguments there remain no obvious reasons not to recognize decompositional semantic structure.

In this section I will argue that decompositional accounts of intuitive, pretheoretic semantic properties and relations, as well as the traditional explanations of reference fixing (properly construed), informal inference and language understanding rejected by FGWP, are in fact preferable to their primitivist alternatives. There are therefore good reasons to accept the decompositional account of language structure. I will also suggest a general approach to psycholinguistic investigation that strikes me as more likely to yield conclusive results on the question of the psychological reality of decompositional representations.

*4.1. Intuitive semantics*

Speakers of natural languages routinely ascribe such properties as *sameness of meaning* (synonymy), *repetition of meaning* (redundancy), *opposition of meaning*

(antonymy), and the like to expressions of their languages. Any competent speaker of English, for example, will intuitively recognize the sameness of meaning of "father" and "male parent", the repetition of meaning of "male father", the opposition of meaning of "male" and "female", and so on. Given possession of these properties by expressions as data, the question becomes how they are best to be accounted for. The decompositionalist typically provides explanations in terms of structured word meanings and their relations to each other. Thus, the synonymy of "father" and "male parent" is due to "father" having a structured meaning *identical* to that of "male parent"; the redundancy of "male father" is due to the meaning of "male" being *contained in* the complex meaning of "father"; the antonymy of "male" and "female" is due to their structured meanings' containment of *incompatible* meanings. Such notions as identity, containment and incompatibility are defined in terms of the geometry of decompositional representations in an optimal semantic theory [23].

The non-decompositionalist, in contrast, must take possession of these intuitive semantic properties to be primitive: it is just a brute fact about language that "male parent" and "father" have the same meaning, that "male father" is redundant, that "male" and "female" are antonymous, etc. Of course, as FGWP point out, any theory that takes meaning seriously will have to count *some* semantic facts as primitive. But surely the preferable theory is the one that (1) minimizes the number of brute facts and (2) does not ignore structure when it exists.

### *4.2. Reference fixing*

In Section 3.1.1, I distinguished reference assignment and reference constraint, and argued that the latter is the true function of definitions in the theory of reference. FGWP suggest that the best bet for a theory *assignment* of reference to primitives is one that focuses on the *causal* aspect of the empiricist account [24]. Let us assume this view, and consider which theory, the decompositional or the primitivist, provides the better explanation of reference *constraint*.

Both theories must explain such facts about reference as, for example, that "father" and "male parent" are necessarily coextensive; and both can appeal to meaning equivalence (synonymy) to do so, claiming that synonymous expressions generally have the same extension as a matter of necessity, and that "father" and "male parent" are synonyms [25]. Where they differ is in their account of this relation. As discussed above, the decompositionalist says that "male parent" expresses a *structural analysis* of the meaning of "parent", and that *that* is why their meaning is the same. The primitivist must say that the relation is primitive: it is, again, just a brute fact about language that the expressions "father" and "male parent" are synonymous; there *is* no explanation of the synonymy. Hence, the explanation of necessary coextensiveness is relatively shallow.

Assuming, then, that the theory with a deeper story to tell, all things equal, is preferable, the decompositional theory of reference constraint is preferable to its primitivist alternative.

*4.3. Informal inference*

FGWP argue that intuitions of informal validity are not intuitions of definitional relations, and suggest that they are either reports of empirical beliefs or intuitions of deductive relations determined by meaning postulates. This is not the place for a discussion of Quinean epistemology; so let us focus on the second suggestion.

The decompositionalist claims that the validity of *some* non-logical inferences can be explained by appeal to the semantic structure of terms. "Bob is a father" implies "Bob is a male" because the meaning of the former contains the meaning of latter (the meaning of "father" contains the meaning of "male"); "The sky is blue" implies "The sky is colored" because the meaning of the former contains the meaning of the latter (the meaning of "blue" contains the meaning of "colored"). Moreover, as noted above, the decompositionalist does not claim that his theory explains *all* valid informal inferences, but only those whose validity depends on the contents of constituent terms. Intuitively valid informal inferences inexplicable in these terms it leaves unexplained.

Not surprisingly, the primitivist must hold that inferential relations that depend on semantic relations among constituent terms are, like those relations themselves, primitive. He will thus have to employ meaning postulates, which record intuitively non-logical inference relations among predicates without displaying lexical semantic structure, in his account of informal inference. A meaning postulate approach will, however, both over- and undergeneralize.

Meaning postulates do not distinguish between intuitively semantic informal inferences and intuitively *non*-semantic informal inferences. For example, the validity of the inference "Bob is blue; therefore, Bob is colored" is given the same "explanation" as the validity of "Bob has two parents; therefore, Bob has fewer than six parents": both depend on a suppressed premise (a meaning postulate) that formalizes an entailment relation between the predicate of the premise and the predicate of the conclusion. Given, however, that there is a way to make theoretical sense of the intuitive differences among such inference types (as in Section 3.2), this explanation overgeneralizes, classing inferences of distinct types together.

Moreover, if in general a level of grammatical structure exists over which inference rules may be defined, to ignore that structure in constructing an account of inference is to *under*generalize. Thus, for example, to account for the validity of the inference from "Bob has a father" to "Someone has a father" in a *sentential* calculus by assigning "$p$" to the first sentence and "$q$" to the second, and adding the postulate "$p \rightarrow q$", is to miss a generalization [namely, that for any $a$ and $F$, $Fa \rightarrow \exists x(Fx)$] [26]. By exposing another level of grammatical structure (subject–predicate structure), the predicate calculus provides a domain for the formalization and application of inference rules. Meaning postulates undergeneralize because they ignore a level of grammatical structure over which inference rules may be defined that subsume such inferences as that from "Bob is a father" to "Bob is a male".

Therefore, since meaning postulate accounts both over- and undergeneralize in their explanation of informal inference, and decompositional accounts don't, the latter are preferable.

### *4.4. Sentence comprehension*

I have argued that FGWP's psycholinguistic procedures are ill-suited to the detection of decompositional structure in mental representations. In closing, I will make some suggestions as to how such structure may be more profitably tested for. These suggestions will be in the spirit of FGWP's proposal, though I consider the relevance of psychological reality to linguistic theory to be an open question.

On the face of it, direct questioning about the semantic properties and relations of expressions would seem the obvious strategy for a chronometric test in FGWP's style. Suppose, for example, that one were to predict on the basis of the proposed semantic analysis of two terms, $\alpha$ and $\beta$, that two expressions containing them, $\Sigma_1$ and $\Sigma_2$, respectively, are redundant. If, further, the proposed analysis of $\alpha$ were more complex than that of $\beta$ [27], then one might predict that if speakers judge $\Sigma_1$ and $\Sigma_2$ to be redundant, there would be some measurable behavioral manifestation of the difference. If this is in fact the case—if speakers judge $\Sigma_1$ and $\Sigma_2$ to be redundant, and there *are* the expected behavioral (chronometric) differences—then one's hypothesis as to (at least) the relative semantic complexity of $\alpha$ and $\beta$ is confirmed (even if one hasn't quite got the analyses right).

Some such procedure seems more likely to produce semantically relevant psychological results than the one exploited by FGWP. It is also more directly relevant to the question of the theoretical utility (and, hence, justification) of decompositional semantic representations, in that it concerns itself with the type of pretheoretic semantic phenomena that are, arguably, the true province of a decompositional theory—synonymy, antonymy, ambiguity, redundancy, analyticity, and the like.

## 5. Conclusion

Since Quine, most analytic philosophy has been done on the assumption that the analytic–synthetic distinction is untenable. Since Fodor, Garrett, Walker and Parkes, a great deal of cognitive science has been done on the assumption that lexical items and their associated conceptual representations are semantically simple. If Katz's arguments and the arguments presented in this paper are sound, philosophy and cognitive science have imposed on themselves limitations of explanatory resources that are both counterproductive and unnecessary. If a path to more adequate explanations in these areas has been cleared, it is surely in everyone's interest that it be taken.

## Acknowledgments

Thanks to Brad Armour-Garb, Jody Azzouni, Ed Becker, Bob Fiengo, Jerry Katz, Arnie Koslow and Steve Schiffer for helpful comments on earlier drafts. Special thanks to Jerry Katz for many happy hours of philosophical conversation.

## Notes

[1] By "semantically structured" I mean "expressing a meaning (sense) that contains other meanings (senses)". Containment should be understood in the Kantian way (this can be cashed out in terms of properties of formal representations in a semantic theory). I will also refer to this kind of structure as "decompositional", and the property of having it as "decompositionality".

[2] FGWP (1980) describe decompositionality as "one of those ideas that hardly anybody [in the cognitive sciences] ever considers giving up" (p. 264). MacNamara and Miller (1989, p. 360) term FGWP (1980) "[t]he most imposing paper, by far, on semantic decomposition", noting that it is "often cited as sounding the death knell for decompositional theories". Smith *et al.* (1984, p. 272) cite it to support their claim that "[a]ll the evidence suggests that the classical [(decompositional)] view is wrong as a general theory of concepts". See also Lucas (1991, p. 258).

[3] An earlier effort by Fodor *et al.* (1981) along the same lines is generally taken to have been successfully criticized by Katz (1981a). See Gergely and Bever (1986). Fodor (1981) also contains an argument against definitions, which I criticize in Pitt (1994).

[4] Katz is, of course, not the only philosopher to resist Quine's conclusions on analyticity (see for a recent example, Boghossian, 1996); but, as far as I know, Katz's arguments are the only ones Quine has explicitly acknowledged as having any real force against him.

[5] Bracketed page number references herein will be to this article.

[6] Representations of sentences need only include "what may be required for the representation of such relations as quantifier binding, operator scope, etc." [p. 311] (i.e. logical form).

[7] A fourth class of arguments, those based upon theories of concept acquisition, is mentioned, but, as far as I can see, there are no arguments on offer based on such theories. FGWP merely emphasize the importance of the simple/complex distinction for traditional theories of concept learning.

[8] An analogous distinction can be made with respect to logical truth. Quine (1986) characterizes logic as the theory of sentences whose truth (or falsity) is *assured by*, though not *a matter of*, their logical structure (p. 48), and decries what he calls "the linguistic theory of logical truth", which has it that logical truths are true "purely by virtue of language" (p. 96). Consistent with this view of logical truth, we may identify assigning and constraining roles of logical structure in fixing truth value. The logical structure of any sentence fixes its truth value, in the assigning sense, relative to a provision of extensions for its referring expressions. What is distinctive of tautologies and contradictions is that their logical structure *constrains* their possible truth values, independently of any assignment of extensions. It would be a confusion to hold that the former role is discharged *by virtue of* the latter role's being discharged—i.e. that reference is assigned because it is constrained. Now, I am not accusing FGWP of an analogous confusion with respect to terms, but, rather, of failing to appreciate one side of the distinction. [This conception also provides for a notion of *analyticity* on which it is not truth by virtue of meaning (as would be claimed by a "linguistic theory of *semantic* truth"), but a distinct type of grammatical structure that constrains the possible truth values of sentences—a type of structure due, in fact, to the semantic structure of constituent terms (see Katz, 1992, for discussion).]

[9] This objection echoes Lewis (1983) and Searle (1974). See Katz (1990, pp. 211–215) for a line of response somewhat different from the one I am pursuing here. (Katz's emphasis is on the interpretive relation between formal representations and the senses they represent. The point I am pressing in this section is not that there are two sorts of interpretation, but only that expressions may bear definitional relations to each other that constrain their reference. Though definitions are typically taken to express intensional structure, they don't have to be: definability does not *entail* intensionalism.)

[10] Indeed, this is the sort of view FGWP themselves consider the most likely option: "The best current hope for [a theory of language and the world] is perhaps to accept that aspect of the Empiricist treatment of primitive terms which claims that the relation between words and their extensions is somehow mediated by causal chains, but to abandon the condition that the relevant

chains are exhaustively specifiable by reference to the behavior of sensory/motor mechanisms" [p. 308]. [This is also the view Fodor (see 1990, Chapter 4) currently endorses.]

[11] *Vide* the endorsement of primitivist semantics on the basis of the failure of "definitional" theories of informal validity, ca. p. 272. See also Fodor (1981, p. 292): "It's ... presupposed that complexity implies definability ...".

[12] See also Prior (1949) for a thorough discussion of determinables and determinates; though Prior does not end up classing arguments such as (5)–(6) as analytic.

[13] I argue that this early proposal is superior to that offered in Katz (1987, p. 228, note 16), on which the differences among color concepts are represented as having conceptual content (in Pitt, 1994, Chapter 5). [Katz (1997) returns to his original account.] I also suggest there that distinguishers might be used in the semantic analysis of natural kind terms. "Horses are animals", for example, might (*pace* Putnam) be analytic, though there is no determinate predicate *P* to be conjoined with "animal" such that "animal and *P*" is analytically equivalent to "horse" ("equine animal", "feline animal", "porcine animal", etc. might be just as bogus as definitions of "horse", "cat", "pig", etc. as "rubine color", "verdine color", "purpine color", etc. would be as definitions of "red", "green", "purple", etc.).

[14] See Gergely and Bever (1986) for a different approach to non-definitional decompositionality, developed in response to FGWP's experimental results. Gergely and Bever's paper is typical of responses to FGWP in the linguistics literature in its exclusive focus on the psycholinguistic argument (see e.g. Jackendoff, 1983; Gonsalves, 1987; Lucas, 1991).

[15] For example, assuming that "father" is defined by "male parent", note that though $(x)[\text{father}(x) \rightarrow \text{male}(x)]$ and $\neg(x)[\text{male}(x) \rightarrow \text{father}(x)]$ (cf.: "$(x)(\text{polygon}(x) \rightarrow \text{plane figure}(x)$" and $\neg(x)(\text{plane figure}(x) \rightarrow \text{polygon}(x))$), MALE is not a superordinate of FATHER (being a parent does not *presuppose* being male). Meaning postulates do not suffice to distinguish hierarchical from non-hierarchical asymmetric entailments. Some other style of representation is needed (see e.g. Jackendoff, 1990; Katz, 1990; Sommers, 1982, for suggestions incorporating non-standard formalisms).

[16] Cf. the papers mentioned in note 14 for other rejoinders to FGWP's psycholinguistic argument.

[17] See e.g. Fiengo (undated), for an argument for the independence of psychological and syntactic theories, and Katz (1981b), Langendoen and Postal (1986), Soames (1985), and Katz and Postal (1991), for arguments for a platonistic foundation for linguistics.

[18] Such a test procedure cannot be validated on causative sentences themselves, since it would be unclear whether a negative result (i.e. no evidence of shift) would indicate no shift, or insensitivity to decompositional shifting. Also, note that the requirement that there be structural differences between surface and semantic levels may be satisfied by non-definitional as well as definitional analyses.

[19] Speakers' relatedness ratings may be used to construct hierarchical representations which closely resemble derived phrase markers.

[20] Levelt explicitly assumes that it is (1970, p. 113).

[21] That is, the second "John" is present at surface structure, but is phonologically unrealized. Note that this will also be true of the gapping cases Levelt mentions in his paper, such as "John eats apples and Peter pears". ("Eats" is present at surface in the second clause, but is phonologically unrealized.) See Fiengo and May (1995, Chapter 4).

[22] In GB syntax, for example, thematic roles are taken to be assigned at D-structure [the "theta criterion"—see Chomsky (1981)—is a constraint on D-structure]; D-structure is not a semantic level of representation.

[23] See e.g. Jackendoff (1972, 1990), Katz (1972, 1990), Sommers (1982).

[24] Though they cite "deep troubles" for it, including "abstract reference, reference to fictions and the like" [p. 309]). Fodor himself has since become somewhat more optimistic (see e.g. Fodor, 1987, 1990, 1994).

[25] Recall that intensionalism is neither necessary nor sufficient for decompositionalism.

[26] This point was made by Katz (1977).

[27] Where relative complexity is, again, defined in terms of the geometry of representations in an optimal semantic theory.

## References


BOGHOSSIAN, P.A. (1996). Analyticity reconsidered. *Noûs, 30*, 360–391.

CHOMSKY, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

CLARK, C. (1993). *Meaning, skepticism, and truth in the immanent naturalism of W.V. Quine*. Doctoral dissertation, New York: CUNY Graduate Center.

FIENGO, R. (undated). On the relation of linguistics to other fields. Unpublished manuscript.

FIENGO, R. & MAY, R. (1995). *Indices and identity*. Cambridge, MA: MIT Press.

FILLMORE, C.J. (1968). The case for case. In E. BACH & R.T. HARMS (Eds) *Universals in linguistic theory*. New York: Holt, Reinhart & Winston.

FODOR, J.A. (1981). The present status of the innateness controversy. In J.A. FODOR (Ed.) *Representations*. Cambridge, MA: MIT Press.

FODOR, J.A. (1987). *Psychosemantics.* Cambridge, MA: MIT Press.

FODOR, J.A. (1990). *A theory of content and other essays*. Cambridge, MA: MIT Press.

FODOR, J.A. (1994). *The elm and the expert.* Cambridge, MA: MIT Press.

FODOR, J.A., GARRETT, M.F., WALKER, E.C.T. & PARKES, C.H. (1980). Against definitions. *Cognition*, *8*, 263–367.

FODOR, J.D., FODOR, J.A. & GARRETT, M.F. (1981). The psychological unreality of semantic representations. In N. BLOCK (Ed.) *Readings in philosophy of psychology*, Vol. 2 (pp. 238–252). Cambridge, MA: Harvard University Press.

GERGELY, G. & BEVER, T. (1986). Related[ness] intuitions and the mental representation of causative verbs in adults and children. *Cognition, 23*, 211–277.

GONSALVES, R.J. (1987). For definitions—a reply. *Cognition, 29*, 73–82.

GRUBER, J.S. (1965). *Studies in lexical relations*. Doctoral dissertation, Cambridge, MA: MIT Press.

HIGGINBOTHAM, J. (1985). On sematics *Linguistics Inquiry*, *16*, 547–593.

JACKENDOFF, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

JACKENDOFF, R. (1983). *Semantics and cognition*. Cambridge, MA: MIT Press.

JACKENDOFF, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

JOHNSON, W.E. (1921). *Logic*. New York: Dover.

KATZ, J.J. (1967). Some remarks on Quine on analyticity. *Journal of Philosophy*, *LXIV*, 36–52.

KATZ, J.J. (1972). *Semantic theory*. New York: Harper & Row.

KATZ, J.J. (1977). The advantage of semantic theory over predicate calculus in the representation of logical form in natural language. *Monist*, *60*, 380–405.

KATZ, J.J. (1981a). The real status of semantic representations. In N. BLOCK (Ed.) *Readings in philosophy of psychology*, Vol. 2 (pp. 253–275). Cambridge, MA: Harvard University Press.

KATZ, J.J. (1981b). *Language and other abstract objects*. Totowa: Rowman & Littlefield.

KATZ, J.J. (1987). Common sense in semantics. In E. LEPORE (Ed.) *New directions in semantics* (pp. 157–233). London: Academic Press,

KATZ, J.J. (1988a). *Cogitations*. New York: Oxford University Press.

KATZ, J.J. (1988b). The refutation of indeterminacy. *Journal of Philosophy*, *LXXXV*, 227–252.

KATZ, J.J. (1990). *The metaphysics of meaning*. Cambridge, MA: MIT Press.

KATZ, J.J. (1992). The new intensionalism. *Mind*, *101*, 1–31.

KATZ, J.J. (1997). Analyticity, necessity and the epistemology of semantics. *Philosophy and Phenomenological Research*, *57*, 1–28.

KATZ, J.J. & POSTAL, P. (1991). Realism vs. conceptualism in linguistics. *Linguistics and Philosophy*, *14*, 515–554.

LANGENDOEN, D.T. & POSTAL, P.M. (1986). *The vastness of natural languages*. Oxford: Basil Blackwell.

LEVELT, W.M. (1970). A scaling approach to the study of syntactic relations. In G.B. FLORES D'ARCAIS & W.J.M. LEVELT (Eds) *Advances in psycholinguistics*. Amsterdam: North Holland.

LEWIS, D. (1983). General semantics. In D. LEWIS (Ed.) *Philosophical papers,* Vol. I (pp. 189–232). New York: Oxford University Press.

LUCAS, M.M. (1991). Semantic representation of meaning—a defense. *Psychological Bulletin*, *110*, 254–263.

MACNAMARA, T.P. & MILLER, D.L. (1989). Attributes of theories of meaning. *Psychological Bulletin*, *106*, 355–376.

PITT, D. (1994). *Decompositional semantics*. Doctoral dissertation, New York: CUNY Graduate Center.

PRIOR, A.N. (1949). Determinables, determinates, and determinants. *Mind*, *58*, 1–20, 178–194.

QUINE, W.V.O. (1966). Carnap and logical truth. In *The ways of paradox*. Cambridge, MA: Harvard University Press.

QUINE, W.V.O. (1967). On a suggestion of Katz. *Journal of Philosophy*, *LXIV*, 52–54.

QUINE, W.V.O. (1980a). Two dogmas of empiricism. In *From a logical point of view*. Cambridge, MA: Harvard University Press.

QUINE, W.V.O. (1980b). The problem of meaning in linguistics. In *From a logical point of view*. Cambridge, MA: Harvard University Press.

QUINE, W.V.O. (1986). *Philosophy of logic*. Cambridge, MA: Harvard University Press.

QUINE, W.V.O. (1990). Comments on Katz. In R. BARRETT & R. GIBSON (Eds) *Perspectives on Quine*. Oxford: Basil Blackwell.

RAPPAPORT, M., LEVIN, B. & LAUGHREN, M. (1993). Levels of representation. In J. PUSTEJOVSKY (Ed.) *Semantics and the lexicon* (pp. 37–54). Dordrecht: Kluwer Academic.

SEARLE, J. (1959). On determinables and resemblance. In *Proceedings of the Aristotelian Society*, *33*, 141–158.

SEARLE, J. (1967). Determinables and determinates. In P. EDWARDS (Ed.) *The encyclopedia of philosophy*, Vol. 2, New York: Macmillan Publishing Co., Inc. & The Free Press, 357–359.

SEARLE, J. (1974). Chomsky's revolution in linguistics. In G. HARMAN (Ed.) *On Noam Chomsky*. New York: Anchor Books.

SMITH, E.E., MEDIN, D.L. & RIPS, L.J. (1984). A psychological approach to concepts: comments on Rey. *Cognition*, *17*, 265–274.

SOAMES, S. (1985). Semantics and psychology. In J.J. KATZ (Ed.) *The philosophy of linguistics* (pp. 204–226). Oxford: Oxford University Press.

SOMMERS, F.T. (1982). *The logic of natural language*. Oxford: Oxford University Press.

Copyright of Philosophical Psychology is the property of Carfax Publishing Company and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.