

From Impossible Words to Conceptual Structure: the Role of Structure and Processes in the Lexicon

KENT JOHNSON

Abstract: The structure of words is often thought to provide important evidence regarding the structure of concepts. At the same time, most contemporary linguists posit a great deal of structure in words. Such a trend makes some atomists about concepts uncomfortable. The details of linguistic methodology undermine several strategies for avoiding positing structure in words. I conclude by arguing that there is insufficient evidence to hold that word-structure bears any interesting relation to the structure of concepts.

Introduction

An important finding of theoretical linguistics is that there are certain broad patterns of words that never appear in a language. For instance, although English contains the transitive verb *break*, as in *Mary broke the desk*, English does not contain a verb expressing the converse relation of *break*, as in **The desk blikked Mary*. The absence of a verb like *blik* does not appear to be an accident, in the way that it is an accident that there is no noun that picks out one's tallest friend's relatives. Instead, the absence of the verb *blik* appears to be due to the more general fact that whenever a transitive verb of English expresses a relation between the doer of an action and the thing that is acted upon — what linguists call the 'agent' and the 'theme' of the verb — the former is the subject of the verb and the latter is the object (e.g., *blacken*, *boil*, *kill*, *shrink*, etc.). Like all generalizations in linguistics, this one is subtle. Since I will use it to motivate the issue, a few comments are in order. First, 'agent' and 'theme' are technical terms of linguistics, and they apply (as does this entire discussion) only to a purely internal conception of language (e.g. Chomsky 1986). These linguistic notions are not the same ones that may be subjected to conceptual analysis in the philosophy of action. Rather, they are categories that our language uses to organize the participants of an event named by a verb. Second, the generalization does not dictate the subject and object of every verb: if your coffee cup resembles mine, for instance, then mine may resemble yours. However, *resemble* does not clearly express that one thing acts upon or does something to another. Our focus will be on the really clear cases, like those listed above (which is not to say there are no unclear cases; cf. Dowty 1989,

I am grateful to John Kulvicki, Robert Matthews, Robert May and two anonymous reviewers for useful discussion and criticism.

Address for correspondence: Department of Logic and Philosophy of Science, 3151 Social Science Plaza A, University of California, Irvine, Irvine, CA 92697-5100

Email: johnsonk@uci.edu

1991). Third, the passive form of verbs, as in *The desk was broken by Mary* is not relevant here, because the passive form is a morphosyntactic alternation of a root verb (e.g. Baker, Johnson and Roberts 1989). The generalization applies only to root verbs, of which there are many like *break* and none like *blik*.

When a candidate word violates a generalization like the one just discussed, it is called an *impossible word*. Here, *impossible* just means impossible relative to the language in question, or perhaps relative to all natural languages. That is, impossible words are ungrammatical words. The linguistics literature contains a variety of explanations why words like *blik* are impossible (e.g., Baker 1988, Grimshaw 1990). More generally, the phenomenon of impossible words has received much attention from widely differing standpoints (e.g., McCawley, 1968; Carter, 1976; Dowty, 1979; Baker, 1988; Grimshaw, 1990; Hale and Keyser, 1993, 1997, 1999; Fodor and Lepore, 1999; Fodor, 1998; Jackendoff, 1990; Zubizarreta, 1987).

The fact that there are impossible words is often taken to have important implications for psychology and the philosophy of mind. This is because it is common for linguists to explain why some words are impossible by appealing to complex structure within actual words. All of the citations above give examples of this last claim, and I will illustrate how it works below. Indeed, the existence of impossible words and the theories that explain them often provide some of the strongest linguistic support for the claim that many of our ordinary words contain complex linguistic structure. The bridge from linguistics to psychology takes the form of an assumption that (*ceteris paribus*) the structure of our words is mirrored in the concepts they express. There are various ways of spelling out this word-concept bridging assumption. For definiteness, I focus only on what I'll call the 'Isomorphism Assumption', which says that the structure of a word is isomorphic to the structure of the concept it expresses (assuming that the word expresses a concept).¹ Many philosophers and linguists hold this view, or something like it. For instance, Jerry Fodor and Ernie Lepore write:

The idea that quotidian, middle-level concepts typically have internal structure — definitional, statistical, or whatever — plays a central role in practically every current approach to cognition. Correspondingly, the idea that words that express quotidian, middle-level concepts have complex representations 'at the semantic level' is recurrent in linguistics; it is the defining thesis of what is often called 'lexical semantics' (Fodor and Lepore 1999, p. 445).

In a related spirit, Jaklin Kornfilt and Nelson Correa describe their theory of word-structure as 'part of the larger system of conceptual structure which underlies

¹ Although there are many interesting issues about the relation between words and concepts that might also be explored — e.g., what does it mean for word-structure to correspond to conceptual structure, what happens if we weaken the relation between words and concepts to a homomorphism, or strengthen it to identity — I will leave these issues aside, since they will not affect the main line of the paper.

human cognitive abilities' (Kornfilt and Correa, 1993, p. 79; cf. also Jackendoff, 1990, esp. ch. 1). Similar views have been echoed in the philosophical and psychological literature (e.g., Peacocke, 1992, p. 3, fn. 2; Fodor, 1998, p. 2; Miller, 1996, pp. 18–19). If such sentiments are correct, then we might be able to learn about our concepts — how many we have, which ones we have, how they are structured, etc. — by looking to the structure of language. Similarly, it might turn out that linguistics alone can justify the claims that some words have definitions and/or that some concepts have possession conditions involving other concepts (e.g., that *break* as in *Mary broke the desk* really does mean *cause to be broken*, and that to possess the relational concept BREAK, one must also possess a certain concept of causation). From here, it is only a short road to establishing the existence of analyticities, synonymies, and many other hotly contested issues from the history of philosophy. The Isomorphism Assumption also explains why people who care about concepts often care about the nature and structure of words. In particular, those who hold that 'most lexical concepts [sc. the concepts expressed by our individual words] have no internal structure' will care how the linguistic theory turns out (Fodor, 1998, pp. 2, 121; cf. Fodor and Lepore, 1992; Margolis and Laurence, 1999). Such researchers will be strongly motivated to show that words have no internal structure. They will want to deny that theories of impossible words can supply evidence that words have internal structure.²

Although the phenomenon of impossible words initially appears to be an obscure topic from a small corner of linguistics, we see that it is potentially quite important for the study of broader psychological issues. However, the empirical and conceptual issues involved in research into the nature of words are both delicate and complicated. In this paper, I sort out a few of these issues. I focus on a number of interpretations of linguistic practice which suggest that despite what linguists say, words contain no interesting structural elements. Some of these suggestions can be found in the literature, and others are novel, but have an intuitive appeal. If words have no interesting structure, then by the Isomorphism Assumption, it would follow that the concepts they express lack structure (or at least the kinds of structure that linguistics is often thought to uncover). Such a result would be a major advance for proponents of conceptual atomism (e.g. Fodor, 1998; cf. Fodor and Lepore, 1992). However, I will argue that none of these suggestions show that words lack structure. Moreover, since these suggestions appear to be the most plausible ones, it looks as though conceptual atomism does not have a correlate in natural language.

² This way of putting things is a little loose. Everyone can agree that there is structure in a word that has no conceptual correlate; for instance, the phonological properties of words are probably not mirrored in our concepts. The issue is rather whether the concept that a word expresses contains all (and only, perhaps) the word's semantic — and semantically relevant syntactic — structure. Since most of what I have to say will concern fairly straightforward examples like whether the transitive verb *break* contains a causal element, the admittedly rough characterizations given here will suffice.

The paper is structured as follows. In section 1, I characterize how a theory of impossible words putatively supports theories about the internal structure of words. In sections 2, 3, and 4, I address three general strategies for avoiding the conclusion that words contain internal structure, despite what the evidence from impossible words suggests. All of these strategies fail, for reasons intimately concerned with the methodology of linguistics. In section 5, I argue that the Isomorphism Assumption is too tendentious and unclear to be of much service to cognitive science. I conclude in section 6.

1. Impossible Words and the Arguments They Figure Into

In this section, I provide a first-pass characterization of how linguistic theories of impossible words might be thought to support the existence of internal structure in many of our words and concepts. The rest of the paper will be based around attempts to undermine this argument. The argument is best illustrated with an example. I will use a simplistic theory of the English causative construction. Before sketching this simplified theory, let me note that the objections to be considered are intended to hold against *any* theory of the lexicon whatsoever. Two of the objections come from Fodor and Lepore, who are adamant about this point: 'We think there are principled objections to this form of argument' (Fodor and Lepore 1999, p. 445), 'impossible-word arguments are infirm *in principle*' (p. 447 emphasis added; cf. also p. 446). So strictly speaking, the present theory of causative verbs will serve as a counterexample to the objections. However, I have deliberately chosen a well-known and traditional theory to illustrate in general how theories of this general type operate. So if your favorite view of the lexicon does not agree on every aspect of the sample theory, you should be able to adjust the examples without altering the main point.³ Thus, I hope to show not merely that the objections in question are unsuccessful in principle but that they are unsuccessful in practice and may be ignored.

The simplified theory I'll use addresses the question why we have verbs like *break* (as in *Mary broke the desk*) but not ones like *blik* (as in **the desk blikked Mary*). To account for this phenomenon, the theory hypothesizes that although the transitive verb *break* appears to be a simple lexical item, it is in fact a complex structure, derived from the intransitive verb *break* (as in *the desk broke*), and another structure that means roughly CAUSE.⁴ Thus, there might be some processes in the

³ In terms of a larger linguistic theory, I will be assuming a very mainstream form of Principles-and-Parameters (Chomsky 1981). However, the general strategy I offer should be amenable to a wider variety of linguistic theories, although some of them may need to fiddle with some the arguments a little. E.g., proponents of Lexical-Functional Grammar will want to alter the argument in section 4.

⁴ Famously, the causal morpheme does not mean the same thing as the English word *cause* (e.g., Fodor, 1970; Comrie, 1985; Pietroski, 1998, 2000, ch. 1). I return to this point later.

lexicon that allow the intransitive verb *break* to combine with this causal prefix to form a complex structure that means roughly X CAUSE Y BE BROKEN. A crucial part of the linguistic construction of this verbal structure is a 'lexicalization' process, whereby CAUSE and BE BROKEN are fused into a single word, the transitive verb *break*.⁵ Thus, the process of lexicalization takes a structure of the form [X CAUSE Y BE BROKEN] as input, and yields a structure of the form [X BREAK Y] as output. Obviously, this process is in part phonological: *cause* is not a syllable of *break*. Below I offer empirical evidence that the lexicalization process also has syntactic-cum-semantic effects as well. In addition to explaining how causative verbs like transitive *break* (and *boil*, *blacken*, *kill*, *shrink*, etc.) work, such a theory also appears to explain why there are no verbs like *blik*. The process for generating transitive verbs from intransitive ones forces the agent (i.e., the subject of CAUSE) to appear as the subject of the resulting complex verb. Since *blik* places the agent in the object position, it is not generated by the process, so the theory correctly predicts the word is not found in English.

As a first pass, a theory of impossible words of the sort just given appears to license something like the following, which I will call the *Impossible Word Argument*:

- (1) According to theory T, there are processes in the lexicon for altering a word's structure, and for combining various structures into a single structure.
- (2) According to T, impossible words like *blik* are not the output of any processes in the lexicon. So T correctly predicts that the impossible words do not exist.
- (3) Thus, T *explains* why impossible words are impossible.
- (4) Thus, if we are justified in accepting T, we are (*ceteris paribus*) justified in supposing that there are processes and structure in the lexicon.
- (5) [The Isomorphism Assumption:] There is structure in a given word if and only if there is corresponding structure in the concept it expresses.
- (6) Thus, If we are justified in accepting T, we are (*ceteris paribus*) justified in supposing that many of our 'lexical' concepts are structured.

Although this version of the Impossible Word Argument doesn't capture every relevant aspect of the dialectic, it nonetheless allows us to pinpoint where the objections are focused. We will see that the first objection is directed at the inference to (3), and the next two concern the inference to (4).

⁵ Many researchers think that transitive *break* is not a single word, but actually contains some phonologically unpronounced extra structure (e.g., Kratzer, 1996; Travis, 2000). Others disagree (e.g., Grimshaw, 1990, ch. 2; Zubizarreta, 1987). I chose the example because it is a very simple and intuitively plausible illustration of the kind of phenomena and explanations that go on in research into the lexicon. There are many additional examples present in the literature; e.g., the case of *shelve* below.

Before turning to the objections, a word about the Impossible Word Argument is in order. I mentioned earlier that impossible words are simply ungrammatical words. So when a linguist claims that a word is impossible for a given language, she does not mean that it is metaphysically impossible for the word to enter an arbitrary (artificial or natural) language. Rather, the claim is that the psychological laws that distinguish the highly restricted range of natural languages from all the logically possible languages also serve to prevent (*ceteris paribus*) certain words from entering (some of) these languages. That is, the linguist claims that some words are ungrammatical in a fashion similar to the way that **Who did Mary kiss Bill* is. In the latter case, our linguistic abilities do not generate a certain structure, although we can override these limitations by a nonlinguistic mental effort. In the case of impossible words, the evidence suggests that our linguistic abilities do not allow certain structures to be in the lexicon (*ceteris paribus*). Under normal circumstances, humans can, and occasionally do, override these restrictions on the lexicon. But that is only to say that the cognitive laws concerning the lexicon — like virtually every other law of psychology and other special sciences — holds only *ceteris paribus*. Thus, the Impossible Word Argument works by suggesting that our linguistic abilities restrict the ways meanings are fitted onto words, and that the best account of these restrictions is in terms of structure and processes in the linguistic lexicon. So to arrive at (4), it is enough to present the outlines of a plausible theory which appeals to structure and processes in the lexicon. In particular, it is *not* necessary to show that some potential word-meanings are in fact incoherent. It's important to stress this point, because it is frequently missed by philosophers. E.g., Fodor and Lepore accuse linguists of 'forever forgetting' that 'what is supposed to make a word impossible is that there is something defective about its meaning — that is, what makes it impossible that *w* should mean "e" is that "e" is not a possible meaning. (Patently, if "e" is not a possible meaning, then one does not need an explanation of why no word can mean it.)' (Fodor and Lepore, 1999, p. 450, cf. p. 452). Linguists do *appear* to 'forever forget' this issue because it's not the issue that counts. I turn now to the three objections to the Impossible Word Argument.

2. Objection One: Any Word Can Be a Primitive

The first objection to the Impossible Word Argument is due to Fodor and Lepore (hereafter FL). The objection concerns the strategy of explaining the impossibility of a word by showing that the lexicon cannot generate it. We saw such an example above with the case of *blik*. This strategy fails to explain why the language does not contain impossible words, FL argue, because such words might have entered the lexicon as primitive expressions. That is, if FL are right, then there is no reason that a speaker could not learn the word *blik* as a primitive transitive verb. ('Oh no, the sandcastle blicked Sue! That was not nice Sue! Sue ruined a beautiful sandcastle!') If a word can enter the lexicon fully-formed, needing no modification by

processes, then the word itself will not violate any rules before, during or after it is inserted into the syntax.⁶ If this line of reasoning is right, then we have not yet explained why certain words are impossible. Why couldn't *blik* just enter the language as an unstructured primitive element, without undergoing any construction process within the lexicon or elsewhere?

As an argument about the methodology of linguistic and psycholinguistic theory, FL's objection contains a rather straightforward flaw. In order to maintain their position, they must assume that our linguistic abilities enable the lexicon to contain any logically possible verb as a primitive expression. But no contemporary linguist or psycholinguist, as far as I know, would make (or has made) this assumption. To claim that the lexicon can contain any logically possible word as a primitive expression is tantamount to leaving the existence of impossible words unexplained, because the assumption predicts that we will not find the broad patterns of non-occurring words that are so characteristic of the present issue. If the lexicon could contain any primitive word whatsoever, it should contain a significant population of words like *blik*, where the agent is in the object position of the verbs. The fact that we don't find any verbs like *blik* in our language would be utterly amazing if FL's assumption were correct. (It is worth observing that FL offer absolutely nothing in the way of a positive theory of why we have no words like *blik* in our language. Without some means to fill such a lacuna, it is hard to evaluate the strength of their counterproposal.)

I want to press this last point further, because it contains a methodological moral that sometimes goes unnoticed. The importance of explaining broad patterns of non-occurring words is significant *even in the presence of known counterexamples* to the pattern. To illustrate this last claim, I used Beth Levin's (1993) lists of English verbs to collect some transitive verbs which clearly express agency, in the sense that something is clearly doing something to something. My small sample yielded 464 such verbs. Although I was only looking for agentive transitive verbs, it turned out that all 464 of the verbs I collected clearly placed the agent of the action in the subject position and the theme in the object position. 464 is an extremely conservative estimate of the total number of agent-subject verbs; the typical speaker's repertoire probably numbers in the thousands.⁷ But in any case there are at least 464

⁶ FL write: 'By definition, impossible-word arguments purport to explain intuitions of the form "There couldn't be a word *w* that means E" by showing that E couldn't be a derivational source for *w*. But, on the face of it, that doesn't show that there couldn't be a word that means E; the most it could show is that if there *is* a word that means E, then it must be primitive' (Fodor and Lepore, 1999, p. 449).

⁷ To be safe, I did not include verbs that might plausibly be derived from some other linguistic source; e.g., *acidify* is probably a morphological compound of the noun *acid* and the derivational morpheme *-ify*. I did however include some verbs that have nominal counterparts. I assumed, for instance, that *perturb* is plausibly underived, and that the nominalization *perturbation* is probably derived from it. The claim that one verb is derived from some linguistic source and that another is underived is, of course, an empirical claim. This is as it should be. The judgments about agentivity were mine, although they were corroborated 100% by a linguistically untrained colleague. There is no reason to expect a larger sample of judgments would produce any significant variance from these judgments.

transitive verbs expressing agency. I will call transitive verbs with the agent in the subject position ‘agent–subject’ verbs. I found no transitive verbs in English that have the agent as the grammatical object — which I will call the ‘agent–object’ verbs — nor do I know of any.⁸ However, let us assume that there is at least one agent–object verb, and to be safe, let us increase this estimate by an order of magnitude, and assume there are 10 of them, so that we now have a total of 474 transitive verbs with agents and themes. Now suppose that nothing influences the order in which agents and themes appeared in these 474 verbs. In other words, assume with FL that any word can enter the lexicon as a primitive. In that case, we might expect that somewhere around half (237) of them would be agent–subject verbs and the other half would be agent–object verbs. What then is the probability of only ten of these 474 verbs being agent–object verbs and the other 464 of them being agent–subject verbs?⁹ If any such verb is equally likely to be an agent–object or agent–subject verb, a bit of elementary probability theory shows that any given subset of these 474 verbs is as likely to be the set of agent–object verbs as any other subset. Thus, the probability that there would be 10 or fewer agent–object verbs can be computed as below:

$$(6) \quad \frac{\sum_{k=0}^{10} \binom{474}{k}}{2^{474}} = \frac{1.47 \times 10^{20}}{4.88 \times 10^{142}} \approx 3 \times 10^{-123}$$

(6) shows that the chance of there being no more than ten agent–object verbs is vastly less than one out of one googol.¹⁰ If there were 100 agent–object verbs, the probability of the null hypothesis would increase to only 2.65×10^{-57} . In order for there to be even a 1% chance that the two kinds of verbs appear with equal probability, there would have to be at least 395 agent–object verbs. In short, unless the ratio of agent–subject verbs to agent–object verbs becomes vastly smaller than it appears to be, the only plausible option is to assume that there is indeed something about the language that encourages agents to appear in the object position. Another way to see this point is to hold fixed our set of verbs and to adjust the probability of a verb’s being agent–subject. Let p be the probability that a given transitive verb expressing agency is agent–subject. We can then ask, how high would p have to be to ensure that there is a 1% chance that at least 464 out of our 474 verbs are agent–subject? This question is answered by solving equation (7), where p is the probability of a verb’s being agent–subject, and $(1 - p)$ is the probability of its being agent–object.

⁸ Cf. the earlier comments on verbs not expressing agency, and the passive voice of verbs.

⁹ I restrict my attention to those verbs that express agents and themes on the assumption that that linguistic feature is determined by the semantic relation the verb expresses. Thus, when one coins a new verb to denote an action performed by one thing on another, one can create only an agent–subject verb or an agent–object verb.

¹⁰ The expression in the numerator is the sum of the binomial coefficients as k ranges from 0 (no agent–object verbs) to 10 (ten agent–object verbs). The binomial coefficient $\binom{n}{k}$ gives the number of size k subsets present in a set of size n . It is shorthand for $\frac{n!}{k!(n-k)!}$. The equation in (6) corresponds to a one-tailed test of a Bernoulli trial of size 474 with 10 successes where the probability of success is .5.

$$(7) \sum_{k=0}^{10} \binom{474}{k} p^{(474-k)} (1-p)^k = .01$$

(7) shows that p must be greater than 95.79%.¹¹ Given this robust resistance toward agent-object verbs, it is a plausible empirical hypothesis that there really are constraints on how meanings are fit onto words. Although the hypothesis is empirical, it will not be easy to cook up an alternative story that not only accounts for the data as well as the present one but also integrates as well with other neighboring linguistic and psychological theories. Consider, for instance, the following alternative theory. 'Perhaps there are more agent-subject verbs because nowadays when we form new verbs, we have a tendency to structure them like other verbs that we already have, and perhaps we started doing this long ago with a few popular agent-subject verbs. Now there is a preponderance of agent-subject verbs, and the scarcity of the other kind is what makes them seem so awkward.' The problem with this story is that it obviously relies on a notion of when two verbs are relevantly similar. The fact that the relevant similarity is expressed in terms of where the agent is located grammatically makes the story look like a causal explanation of the linguistic generalization, not an alternative to it. Moreover, if this were the end of the alternative story, it would predict that other cultures should develop in which preference was given to agent-object verbs. But this particular linguistic generalization has been investigated in a wide variety of diverse languages, and it appears that there are no such linguistic cultures (cf. Baker, 1988, 1996 for detailed discussion of many typologically distinct languages).¹²

In short, not just any word can be part of the lexicon, *ceteris paribus*. Of course, under highly abnormal circumstances, as when a mad dictator holds a gun to your head, the *ceteris paribus* clause here would be violated. But such thought experiments are not counterexamples; they are observations that the empirical evidence — word usage — is the effect of multiple cognitive abilities. This is typical: it is a primary task of virtually every empirical science to separate out relevant effects of an underlying process from the extraneous 'noise'.¹³ The moral here is that generalizations about the lexicon are statistically significant regularities, so individual counterexamples to them don't necessarily undermine them.

¹¹ For even odds that 464 out of 474 verbs would be agent-subject, we would need $p > .9775$, and for a 90% chance of this result, we would need $p > .9851$.

¹² Two points: (i) The subject of a clause is not defined by its word-order. The notion of a subject is a relatively technical term of linguistics. Within the framework I adopt, the notion of a subject is defined purely configurationally, as e.g., the specifier of a VP or as the specifier of some kind of functional projection associated with the verb. (ii) A complete defense of the claim that there are no languages with agent-object verbs would require a careful study of (inter alia) the ergative languages (cf. e.g., Bittner and Hale, 1996). As I remarked at the beginning, the objections I am considering are supposed to hold in principle, so I am happy to allow my arguments to rest on some empirical assumptions.

¹³ Indeed, this is probably the most commonly used property of statistical analysis; any introductory textbook will contain a range of well-known methods — regression analysis, ANOVA, ANCOVA, etc. — all of which are designed to separate out the effects of a variable from other unexplained sources of variance.

To conclude this section, let us consider normal theorizing about the lexicon. In contrast to FL's proposal, linguists and psycholinguists commonly posit severe constraints on what can enter the lexicon. They often assume that the lexicon contains (i) processes for generating new words out of already existing elements, and (ii) processes for fitting a meaning onto a word, after the meaning has been extracted from the perceived environment. The resulting theories make strong empirical predictions, many of which have been tested and confirmed in experimental settings (e.g., Grimshaw, 1997; Gleitman, 1990; Pinker, 1989, 1999; Gropen *et al.*, 1991, 1991a; Bloom, 2000). The success of the resulting theories stands in direct conflict with FL's assumption. For example, the processes specified in (ii) entail that meanings are fit onto words in certain ways, and not in others. By insisting that a word–meaning be fit onto a word in only a limited number of ways, such theories can predict that words like *blik* do not occur, and explains why they don't (if the theory is correct). So the Impossible Word Argument given in (1)–(4) can be maintained by explicitly adding the background assumption:

- (1.5) If a word is not the output of a process in T, then T predicts that the word is impossible.

On our current assumptions, (1.5) is analytic (or virtually so). After all, in our list of processes, we are including those processes that fit a meaning onto a word, so in that sense every word, even the primitive expressions, are the output of one process or another.

3. Objection Two: Lexicalization is Theoretically Otiose

We have seen that there is good reason to hold that (*ceteris paribus*) not just any candidate word can be contained in the lexicon. Rather, it is much more likely that there are restrictions on the structure of the primitive words of the language. But possessing a word as a primitive element of the lexicon is only one way to have a word in your repertoire. Another potential way is to possess some other words and to have a method for combining them into a structurally complex word. We have seen a simplified example of this in the theory that says that *cause* and *be broken* combine to form the transitive verb *break*. But there are many other such examples. The adjective *black* combines with the morpheme *-en* to form the transitive verb *blacken*, as in *Ted blackened the fish*. An example I will use in what follows is the well-studied case of so-called 'denominal' verbs, i.e., verbs that are derived from nouns. For instance, the transitive verb *shelve* — as in *David shelved the books* — appears to be derived from the noun *shelf*. (The claim that a verb is derived from a noun, and not the other way around, is of course a theoretical claim. For justification of this claim, cf. the references cited below.) Kenneth Hale and Samuel Jay Keyser (HK) have suggested that the verb *shelve* is the result of the lexicalization¹⁴ of a complex syntactic structure,

¹⁴ HK refer to this process as 'conflation'.

roughly the structure of the expression *put on a shelf* (e.g., Hale and Keyser, 1993, 1997, 1999, 2003). Recall that lexicalization is the process that takes certain complex linguistic structures — like *put on a shelf* — as input and derives individual words — like *shelve* — as output. More precisely, the outputs of lexicalization act as individual words on the ‘surface’ syntax of the sentence.

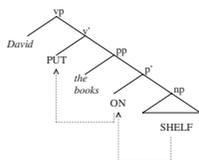
A unique aspect of HK’s theory is that the lexicon is assumed to contain the same sorts of syntactic structures that are found in the overt syntax. That is, they hold that words share the same sort of syntactic structure that sentences do. If this view is on the right track, then we should not find words that have a meaning and/or structure that would violate syntactic rules found in the overt syntax of natural language. If we do not find such words, then that is evidence for their theory. Details aside, HK argue that we do not find such words. For instance, we do not find words like **shelve*, as in (8)

(8) **David shelved the books on.*

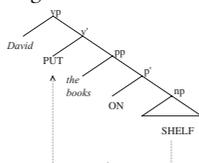
The verb **shelve* in (8) has roughly the meaning that David put the books somewhere with respect to a shelf, and the respect in question is given by the preposition *on* (as opposed to *near* or *under* the shelf). The non-existence of words like **shelve* in (8) is exactly what we would expect if HK’s theory were correct. According to this theory, **shelve* has lexicalized a verb (roughly *put*) and a noun (*shelf*— the indefinite article *a* is not important here), but not the preposition *on* that modifies *shelf*. The lexicalization of all but the preposition would require a structure in the lexicon that violates independently known laws of syntax.¹⁵ Thus, the non-existence of verbs like **shelve* suggests that there are structures and processes in the lexicon.

FL offer a brief argument against this use of impossible words. They can be read as conceding for the sake of argument that the present theory really does provide a useful and important explanation of why some words are impossible (i.e., they can be

¹⁵ For those who are interested in the details, HK’s proposal is that *David shelve the books* has the structure:



Here we see that there is head-movement from *shelf* to *on* to *put*. In contrast, the ungrammatical **David shelve the books on* has the structure:



Here, the head movement omits ON, violating the Head Movement Constraint, which has been independently motivated as a feature of grammar (Travis, 1984; Baker, 1988).

understood as conceding that (1)–(3) are true). But for all that, FL argue, the theory still does not supply evidence that there are processes and structure in the lexicon. Their argument is as follows:

the assumption that lexicalization respects independent constraints on derivations is supposed to explain why intuitively impossible words *are* impossible; and the fact that *if lexicalization obeys the independent constraints on derivations, then it explains why impossible words are impossible* was supposed to be the evidence that there is such a process. [But] if [**shelve*] is ill formed. . .because there is a mandatory constraint that [the lexical structure of **shelve*] fails to meet — then patently [**shelve*] is ill formed whether or not there is a lexicalization transformation. . . .But if [**shelve*] would be ill formed whether or not there is a grammatical process of lexicalization, how could the fact that it is ill formed be evidence for such a process? (FL, 1999, pp. 451–2).

FL's point is this. Suppose it's true that **shelve* can be formed only by giving the lexicalization process an ungrammatical structure as input. Even so, it's still unclear why there should be a process of lexicalization at all, because it's the ungrammaticality of the *input* that explains the impossibility of **shelve*. The process of lexicalization does no theoretical work at all. Thus, even if HK's general story is correct, we still have no reason for positing a mechanism for turning structures into words.

This argument is flawed. Without a process of lexicalization, there will be no way for any complex lexical structures to be formed into words in the first place. According to HK's theory, **shelve* is ungrammatical because below the surface it is a complex ungrammatical structure, and there is a law of language that says that words, just like sentences, must not be ungrammatical structures. But if the process of lexicalization were absent from this picture, there would be no linguistic connection between the verb **shelve* and the ungrammatical structure **put a shelf* from which it was derived. Since the only way the present theory can predict the ungrammaticality of **shelve* is by associating it via lexicalization with an ungrammatical structure, the theory minus lexicalization would be utterly powerless to explain this type of impossible word. Thus, contrary to what FL say, lexicalization is not otiose, but rather is a crucial part of the theory.

Although FL's argument ends here, it may be possible to strengthen it. After all, lexicalization is needed only to relate words to structures in the right way. Perhaps there is some other way to relate them without 'deriving' words from structures as lexicalization does. For instance, we might suppose that all words are unstructured, but that there are nomic relations, underwritten by high-level psychological laws, that hold between various linguistic structures and potential words in the lexicon. For instance, there might be some laws allowing linguistically primitive meanings to be fitted onto unstructured words only if certain correlated structures are grammatical. Here the correlation between words and structures is given by the nomic relation. So the law in question might correlate the possibility of *shelve* with the grammaticality of *put on a shelf*, and the impossibility of **shelve* as

in (8) with the ungrammaticality of **put a shelf*. Since only the first of these expressions is grammatical, only the first word is possible.

Initially, this 'nomic connection' theory looks tempting. Once the relevant laws are specified, the theory makes many predictions regarding impossible words, all the while allowing words to be semantically and syntactically unstructured. Moreover, the theory is flexible: it can make the same predictions as any rival linguistic theory. In fact, the linguistic details of the nomic connection theory can be taken directly from a linguistic theory. Indeed, the nomic connection theory is not a linguistic theory *per se*, but a theory of how linguistic abilities are realized in human beings. The theory produces lawlike statements to the effect that if structure X and word Y have the same meaning (or are otherwise somehow appropriately connected), then Y is a possible word if and only if X is a legitimate structure. No further claim, such as that X *is* the structure Y need be presumed.

If the nomic connection theory is viable, we can use it instead of lexicalization, and thus deny the inference from (1)–(3) to (4) of the Impossible Word Argument. Thus, we needn't concede that there are structures or processes in the lexicon. Unfortunately, the nomic connection theory may not be a panacea. For one thing, not everyone who would like an alternative to the lexicalization process could accept the nomic connection theory. This is because many of the generalizations that the nomic relations will need to explain appear to be stateable only with the help of the technical vocabulary of thematic roles and other sub-lexical semantic constituents (e.g., the notions of an agent or a theme, or of a specialized notion of causation or affectedness, etc.). But some outspoken opponents of structure and processes in the lexicon are skeptical about the use of these other linguistic notions (e.g., FL, 1998, p. 275–278; Fodor, 1998, pp. 57–64). Thus, they could not allow these notions to be part of laws of linguistic psychology without recanting some of their other views. Moreover, the nomic connection theory is *ad hoc*. It's hard to see any reason why these nomic connections should exist. Why should word X be impossible if structure Y is ungrammatical? Why shouldn't X's status depend on the structure Z? Why shouldn't X be possible if and only if Y is ungrammatical? It's hard to see why anyone would adopt the nomic connection theory unless she wanted to save some favored theory of word-structure.

A little linguistic work shows that the nomic connection theory has further problems. Despite appearances, it is not equivalent to more typical linguistic theories that posit structure and processes in the lexicon. I will offer two illustrations of this claim. In the first case, I compare the nomic connection theory to HK's view, discussed above. The strategy I use to show the superiority of the latter view generalizes to a large number of theories, and if the nomic connection view is supposed to be an all-purpose replacement for such theories, it will need to be improved.

Notice that the nomic connection view adds (at least) one new primitive element to the overall theory of human linguistic abilities, namely a lawlike relation holding between structures and words. In contrast, HK's theory works by *reusing* mechanisms that have already been independently established within the linguistic theory. In particular, their process of lexicalization is really the process of 'incorporation', which is well established in the syntax literature (e.g., Baker, 1988, 1997). As originally

introduced, incorporation is a syntactic process occurring when multiple words or morphemes combine to form larger, more complex words. Its syntactic effects are most clearly seen in languages other than English, such as the Bantu language Chichewa. Consider, for instance, the following two Chichewa sentences both of which express that the girl made the waterpot fall. (The examples are drawn from Baker 1988.)

- (10) a. Mtsikana a-na-chit-its-a kuti mtsuko u-gw-e
 girl do cause that waterpot fall
 b. Mtsikana a-na-gw-ets-a mtsuko
 girl fall cause waterpot

In (10a), there are two separate verbs, *its* ('cause') and *gw* ('fall'). In (10b), however, *gw* has been incorporated into the verbal structure for *its*. (The extra material surrounding *its* — *ana-* and *-a* — is inflectional morphology, similar in kind to the English past tense morpheme *-ed*, as in *kicked* or *kissed*. Also, *its* changes to *ets* for irrelevant phonological reasons.) Crucially, when *gw* is incorporated into *its*, the result is a single complex verbal structure that quite literally contains both verbal elements as proper parts. There is, in particular, no reason to suppose that since *its* and *gw* can incorporate, there must also be a primitive verbal word *anagwetsa*, related by some kind of nomic connection to the complex syntactic structure that incorporates *gw* into *its*. So when incorporation applies, the constituents of the input are contained as proper parts of the output, contrary to the way that the nomic connection theory operates. Thus, when the process of lexicalization is identified with incorporation, as HK hypothesize, a single, independently motivated process is responsible for compressing various structures into single words.¹⁶ Moreover, by identifying lexicalization with incorporation, HK's theory thereby predicts that lexicalization will share all of the limitations inherent in the process of incorporation. This prediction is confirmed by the fact that incorporation *cannot* take a structure like *put on a shelf* as input and yield **put-a-shelf on* as output. Thus, the identification of lexicalization and incorporation not only produces a theory of the lexicon whereby structures are squished into words by means of an independently motivated syntactic mechanism, but incorporation is also able to explain why certain relevant structures are ungrammatical. The fact that one independently motivated mechanism does these two further jobs in the lexicon is surely evidence in favor of the overall theory.

On the other hand, the nomic connection theory associates the structures that a linguistic theory produces with individual, unstructured words. In the case in (10),

¹⁶ Two comments. First, this component of HK's theory has received further confirmation as part of the design structure of human languages. Lisa Travis (Travis, 2000) has shown that this general framework is capable of accounting for a number of striking features of Western Polynesian languages like Malagasy and Tagalog. Secondly, for various reasons, HK hold that the verbs and prepositions used in their discussions are not the ordinary ones that are phonologically realized, but are similar lexical ones with no phonological properties at all. Their assumption is in line with much current research into other languages, where similar morphemes with phonological content are common (e.g., Travis, 2000; Baker, 1988, 1997).

the nomic connection theory has absolutely no work to do. Incorporation has already created a structure that behaves like an individual word, so the nomic connection is otiose. Since incorporation gives us an item that acts like a word, why should we insist that the item in question is unacceptable and must be replaced with something else that acts like a word? By Occam's Razor, the nomic connection view is less desirable than the more economical theory that lets incorporation do the work directly.

The second illustration of the difference between the nomic connection theory and standard linguistic theory concerns the former's focus on explaining impossible words. Although the nomic connection theory can (in an extensional sense) account for impossible words, it is unclear how it could also deal with other phenomena that motivate linguists to posit structure in words. For instance, natural languages appear to be sensitive to whether the meaning of a transitive verb expresses that the subject somehow created or obtained the object of the verb (cf. e.g., Pinker, 1989; Levin, 1993). Notice that both *I baked a cake for Marsha* and *I brushed the horses for Marsha* are perfectly normal. However, only the first of these sentences has a variant in the 'double-object construction': *I baked Marsha a cake* is normal, but **I brushed Marsha the horses* is not. A common explanation for this distribution is that only *bake* is a verb of creation. Since the meaning of *bake* expresses a kind of creation of the verb's object, it can appear in the double object construction. However, since *brush* cannot, it cannot so appear (e.g., Pinker, 1989). If this analysis is correct, a linguistic theory will need to appeal to the fact that some verbs express that their objects are created (or obtained). Thus, verbs will need to carry information about some of their semantic properties. By itself, the nomic connection view cannot account for this, since words are supposed to be unstructured entities. While this argument is clearly non-demonstrative, it shows the sort of difficulties and complications involved in maintaining a nomic connection view (cf. Fodor, 1998, ch. 3, esp. pp. 57–64 for some relevant discussion).

In sum, the process of lexicalization cannot just be dismissed, as FL suggest. It plays a crucial role in any theory that allows word-formation processes to occur. Attempting to avoid lexicalization by appealing to nomic connections is *ad hoc*. The nomic connection view packs unwarranted complexity into a linguistic theory, and in return it fails to do much of the explanatory work that a more typical theory of the lexicon does with comparative ease. On balance, then, it appears that lexicalization is here to stay.

4. Objection Three: Put 'Everything In the Syntax' (EIS)

In this section, I consider one final sort of attempt to undermine the need for structure and processes in the lexicon. According to this approach, the lexicon is merely a storehouse of primitive elements, and all the structures and processes posited in the lexicon are simply moved 'upstairs' into the generative component of the grammar, i.e., into the syntax (and combinatorial semantics and phonology)

proper. I will call this view the ‘Everything in the syntax’ proposal, or ‘EIS’ for short. According to EIS, although the lexicon contains simple elements like *cause* and intransitive *break* (assuming these elements are indeed primitive), a word like transitive *break* is formed only after *cause* and intransitive *break* are entered into the syntax of the sentence. Once these elements are in a syntactic structure, they are then organized into transitive *break* in the syntax by the very processes that were thought to occur in the lexicon. Thus, although the lexicon contains a causal predicate and intransitive *break* it does not contain transitive *break*. (As is well known, the causal predicate appealed to here, along with other such linguistic notions as ‘agent’ and ‘theme’, are theoretical terms that do not — and need not — correspond exactly with the intuitive meanings of the corresponding English words. They are justified insofar as they play a role in the theory of language. In particular, their explanatory keep is earned by organizing how we — at least through our language — conceive of events, their constituents, and certain relations holding between them.)

Initially, EIS looks very tempting. EIS employs all and only the mechanisms of a successful linguistic theory, unlike the views discussed earlier. (Of course, EIS will not be available to the likes of Fodor and Lepore, who challenge the viability of the very mechanisms used in linguistic theorizing; e.g., FL, 1999; Fodor 1998, pp. 58ff. However, even they can use EIS with any linguistic theory that does not contain the mechanisms they reject.) Furthermore, EIS produces a simple and restrictive overall theory, since it limits structure and processes to occurring only in the syntax.

Ultimately, the main problem with an EIS approach is that it only rearranges some definitions, leaving all the substantive issues unresolved. EIS avoids the Impossible Word Argument (given in (1)–(4)) by re-defining the lexicon so that all the problematic structures are to be found in the syntax. But the problematic structures were the primary issue in the first place. We began by asking whether transitive *break* is structurally complex. EIS concedes that it is, but insists that *break* is actually a *syntactically* complex structure composed of several primitive expressions, which individually mean something like *cause*, and *be broken*.¹⁷ EIS’s maneuver of putting all the structure into the syntax does *nothing* to reduce the number of semantic primitives contained in transitive *break*. In fact, EIS agrees with the Impossible Word Argument in spirit, since it also maintains that items like *break*, which at least appear to be individual words, are actually complex phrasal constructions.

As another way to view this problem, consider whether EIS can offer any new help for atomists about concepts, such as Fodor. It appears that it can’t. On the one hand, if by the term *word*, we mean an element of the lexicon, EIS does allow us to

¹⁷ In the particular case of causatives, EIS is actually in line with much contemporary thinking (e.g., Hale and Keyser, 1987; Kratzer, 1996; Travis, 2000). As I have discussed, though, we are using a simplified theory of causatives to make a methodological point that applies to many further cases which would be too complex to easily present.

conclude that all words are unstructured, and so by the Isomorphism Assumption, we can conclude that all concepts expressed by words are unstructured, too. But this a hollow victory, because now we need to decide whether something that looks like a word is indeed a word and not a syntactically complex expression, like transitive *break*. So to maintain that the concept corresponding to transitive *break* lacks structure, the atomist will have to show that *break* is a simple word, and this claim is no longer trivial.¹⁸

Although EIS is little help to conceptual atomists, it might still be useful. Atomism aside, is EIS a good way to think about the organization of language? At present, this is a difficult question to answer decisively, because the relation between morphology and syntax is currently the subject of much debate (e.g., Marantz, 1997; Halle and Marantz, 1993). However, I think it is possible to display a general *type* of reasoning which leads to the denial of EIS. The argument I have in mind works by showing what it takes to develop a serious EIS theory of language. Although some of the examples used below are controversial, they still display a style of reasoning that is very common in linguistics. This strategy, I suggest, leads one to accept a *de facto* distinction between the lexicon and the syntax.

In order to develop EIS, we must notice that by itself, EIS does not recover every explanatory advantage of more standard linguistic theories. Let us return to HK's example of the verb *shelve*, as in *David shelved the books* (cf. Hale and Keyser, 1993, 1999, 2003; Fodor and Lepore, 1999). As we saw, HK treat words like *shelve* as having the same structure as larger syntactic expressions. In particular, they treat *shelve* as having all the internal structure of the verb phrase *put on a shelf*. Their theory also explains why there is no corresponding verb **shelve*, as in (8). Details aside, their explanation works by assuming that *on a shelf* is lexicalized with *put* within the lexicon, so that only a verb appears in the syntax. But according to EIS, we could just as well posit the entire phrase *put on a shelf* in the syntax, instead of buried down in the lexicon. However, the EIS proposal cannot yet be complete. For it is well-known that in the syntax, there cannot be multiple independent fillers of the same thematic position (e.g., Chomsky, 1981, 1986 calls this the Theta-criterion). For example, if Julie opened the box using both a hammer and a saw, we would not express this by saying **Julie opened the box with a hammer with a saw*; rather, we would use something like *Julie opened the box with a hammer and (with) a saw*. In the present case, this generalization entails that there cannot be two separate prepositional phrases both of which specify where the book was put, as we see in (14):

- (14) a. David shelved the files on the bookcase;
 b. **David put the files on a shelf on the bookcase.*

¹⁸ Actually, this claim is never trivial, even when one rejects EIS. The point in the text is that the claim becomes hard to justify in precisely those cases that EIS was designed to solve.

(14b) is ungrammatical when both prepositional phrases modify the verb, yielding the meaning that David put the files on a shelf and on the bookcase. (The reading where *on the bookcase* modifies *shelf* is irrelevant here.) According to EIS so far, sentences in (14) should have the same structure. So EIS will need to be augmented to get the facts about (14) right. No doubt this can be done. We simply require that lexicalization operates so that *on a shelf* is absorbed into *put* in such a way that the location of the action of the verb remains unspecified. In other words, the process of lexicalization will solve the problem *as long as it works just as though it was a process of word-formation, taking place in the lexicon before the word is inserted into the syntax*. Thus, lexicalization does its job in the syntax by operating before any other syntactic mechanisms do.

Other putatively lexical processes would need to be handled similarly. For instance, the mechanism that allows some verbs to have a middle form is often thought to occur in the lexicon. We have, e.g., transitive verbs like *sink*, as in *The captain sunk the fiberglass boats*; but we also have intransitive middle verbs like *sink* as in *Fiberglass boats sink easily*. Middle-formation is often thought to occur in the lexicon (at least for English) because it seems to behave differently from other syntactic processes. E.g., when *sink* is a middle, the subject (= agent) position of the verb is no longer present. The absence of the agent of *sink* is robust, insofar as agency cannot be captured by a *by*- phrase, nor can *sink* supply a subject of a purpose clause:

- (16) a. *The fiberglass boats sink easily by crooks.
 b. *The fiberglass boats sink easily to collect the insurance.

(16b) is ungrammatical on the relevant reading 'It is easy for one to sink the fiberglass boats in order that one may collect the insurance on them'. Moreover, speakers experience difficulty with (16b) at exactly the point when they begin to interpret *to collect the insurance* as a purpose clause in need of a subject somewhere in the sentence (e.g. Tanenhaus *et al.*, 1993). Contrast this behavior with the passivization of verbs, which we may assume is a syntactic process. When *sink* is passivized, the subject position of this verb appears to remain in the clause, at least in the sense that it can support *by*- phrases and purpose clauses:

- (17) a. The fiberglass boats were sunk by crooks.
 b. The fiberglass boats were sunk to collect the insurance.

(17b) means that the fiberglass boats were sunk by X in order that X could collect the insurance. Thus, although the subject position of *sink* is not overtly at hand, it still retains some residual presence (cf. Baker, Johnson and Roberts, 1989). We can see that the subject position of *sink* supplies the subject of the purpose clause by noting that (17b) cannot mean that the fiberglass boats were sunk by X in order that Y could collect the insurance on them, where $X \neq Y$. So for EIS, there will again be two sorts of syntactic processes: when transitive *sink* is turned into a

passive, the thematic position for the subject remains, whereas when *sink* is turned into a middle, that position disappears, just as if it were never present in the syntax in the first place.¹⁹

In general, the processes that EIS moves from the lexicon into the syntax are distinctive in that they operate ‘locally’: the total contribution of the output of the process is always located around the ‘visible’ (i.e., phonologically pronounced) output of the process, leaving little evidence of any deleted material. The cases of lexicalizing *put on a shelf* into *shelve* and forming the middle of *sink* are examples of this locality. Many syntactic processes are not local. Consider how an English question word like *who* or *what* appears at the front of a sentence, as in *Who did Peter know Mary gave some candy to?* (cf. *Peter knew Mary gave some candy to John*). Here we see that when *who* appears at the front of the clause, it also produces a change in the location of the past tense morpheme, leaving it at the front, attached to the dummy verb *do* (cf. **Who Mary kicked?*). Furthermore, when a question word is ‘moved’ out of an embedded sentence, there must be an unoccupied position at the front of each clause between the *wh*-word and the position it is associated with.²⁰ If that position is already occupied by another question word, ungrammaticality results: **Who did Peter know what Mary gave to?*

So EIS will probably need to establish a distinction between local and non-local syntactic processes. Further research has shown that local processes also tend to share other features, such as occurring before and independently of any non-local processes. By the time EIS is done with all the rearranging, redefining, and reclassifying, we will be left with a rich and bifurcated theory of syntax. It would be natural for linguists to simply state that one of these two parts is to be called the ‘lexicon’, and the other is to be called the ‘syntax’. Such a maneuver would be natural, because the primary evidence linguists use for determining whether a given phenomenon is (what they call) ‘lexical’ or not is whether the process is local, in the sense described above. Thus, in ordinary linguistic practice, the fact that that a process is local is (*ceteris paribus*) sufficient for declaring the process to be lexical. In short, EIS does not merely import

¹⁹ On a more theoretical note, some linguists may have problems with the conflict between EIS and the Projection Principle. The Projection Principle says that the selectional requirements of a lexical item (including whether it is a transitive or intransitive verb) are present at every level of syntactic representation (e.g., Chomsky, 1981, pp. 29–38, 2000, 2001; Baker, 1988, pp. 49–51). If middle formation is a syntactic process, there should be a level of syntax at which verbs like *sink* and *cut* are represented as transitive verbs, with their subject position being eliminated as per (16). But if this is right, then it looks like the Projection Principle will be false. Of course, an EIS theory can simply restrict the range of the Projection Principle, so that such problems disappear. However, in addition to its increasingly ad hoc appearance, such a move conflicts with the view that the job of syntax is to only organize the various elements of language into some form suitable for phonological and semantic interpretation, and not to alter the fundamental structure of clauses determined by the various lexical items. This view is common in current Minimalist thinking (e.g., Chomsky, 2000, 2001).

²⁰ Strictly speaking, this is not quite right, as can be seen by the counterexample of *Who do you think that the man who knows where the treasure is will befriend?* However, the above crude characterization will suffice for now.

processes from the lexicon into the syntax, it also imports the distinction between the lexicon and the syntax into the syntax. Thus, EIS is merely a notational preference for calling linguistic processes syntactic.

5. The Isomorphism Assumption

In the bulk of this paper, we have seen that there is little reason to question steps (1)–(4) of the Impossible Word Argument. Does this spell the end of conceptual atomism? It does only if the Isomorphism Assumption, which says that word-structure mirrors concept-structure, is true. Initially this assumption looks pretty plausible. If words express concepts and some words are composed of smaller semantic units, then it is plausible that the concepts expressed by the words are themselves composed of these smaller units as well. In the syntactic realm, this looks right: it's doubtful we have a simple concept expressed by e.g., *likes brown horses*. So shouldn't this general idea hold for words too?

This is a difficult question to answer. Relatively little is currently known about the structure and individuation of concepts. Worse yet, in the present case, we cannot address this issue by looking to the structure and individuation of linguistic expressions. Nonetheless, something is surely right about the Isomorphism Assumption: no doubt there is some useful sense in which our words express concepts, and that at least some structurally complex words tell us something important about the concepts they express. But at the same time, there is much about the Isomorphism Assumption that we can criticize. For one thing, it presupposes a prior individuation of words. That is, the content of the assumption is unclear in the absence of a theory of what counts as a word (and similarly for concepts). But as we have seen, saying what counts as a word is a difficult issue.

By adopting the Isomorphism Assumption, conceptual atomists like Fodor face a dilemma concerning the definition of *word*. If transitive *break* counts as a word, then (given current linguistic theory) there can be structurally complex words, and so by the Isomorphism Assumption many ordinary words express complex concepts. But if *break* is not a word, then many items that look, sound, act, etc., like words are in fact syntactically complex structures. But unless we can find some further way of identifying words, conceptual atomism will be empirically useless. If we can't identify the words, then we have no way of discerning whether a given word is conceptually simple, so atomism loses its content.²¹

²¹ This second horn of the dilemma also makes conceptual atomism logically equivalent to the Isomorphism Assumption, if one adopts EIS, because then the lexicon contains no structurally complex items. One might have thought that the theory that 'most lexical concepts [sc. the concepts expressed by our individual words] have no internal structure' expressed more than just the claim that a word's structure mirrors the structure of the concept it expresses (Fodor, 1998, pp. 2, 121; cf. Fodor and Lepore, 1992; Laurence and Margolis, 1999).

Despite these worries, is the Isomorphism Assumption nonetheless plausible, given what we currently know? The linguistic and psychological literature suggest that this is an unresolved matter. On the one hand, a concept could have structure that is not present in the corresponding word. Suppose for instance, that concepts turn out to be the kind of highly structured entities that statistically based theories of meaning (e.g., prototype, stereotype, and exemplar theories) postulate (e.g., Medin, 1989; Smith, 1995). Much of these concepts' structures will not be linguistically relevant, and indeed may simply not be part of the structure of language. If a bit of conceptual structure has no syntactic, morphological or phonological effects, and no compositional effects on the way the concept is used as a constituent of a larger expression's meaning, then that is evidence that the structure is not part of language. So there could be *lots* of structure in our concepts that is not mirrored at all in the structure of our words.

On the other hand, words may well contain linguistic structure that is not present in the corresponding concept. In one sense this is almost trivially true, since words have phonological structure, and concepts presumably do not. But the linguistic lexicon could contain interpretively relevant information that is restricted to the language processing systems, and which serves no purpose in a psychological theory of concepts. For instance, we have seen that a verb's thematic structure is an important grammatical property that is determined by the semantic properties of the verb's meaning. But this thematic structure needn't be present in the structure of the corresponding concepts. Perhaps thematic structure is restricted to the language faculty. For example, we saw above that middle-formation hides the agent of the verb from the syntax. In order for the process of middle formation to work, it must be able to identify the agent in a verb like transitive *sink*, or perhaps the agent in a concept like X *SINK* Y. (Middle formation is much harder for verbs in which no agentivity is present; e.g., **Peter resembles easily*.) Further research has suggested many reasons for supposing that agentivity is somehow part of the linguistic structure of a verb like *sink*. But a linguist who posits this structural property in *sink* needn't also claim that the concept of sinking contains a similar bit of structure. The concept of sinking may be an unstructured unit, even though our linguistic abilities recognizes some linguistically relevant properties (e.g., agency) in it. This additional thematic structure may have been encoded in the lexical entry for *sink* when this verb was acquired, when the concept of sinking was linked to the verb. That is, as part of how we fit concepts onto words, our linguistic ability may screen the concept for various properties. If it has some of these properties, this information may be recorded as part of the structure of the word, and this structure may serve to partially define the word's grammatical role in the language. (The total interaction between syntax and semantics at the acquisition stage is bound to be substantially more complex than this, however; e.g., Gleitman, 1990.) The Isomorphism Assumption could be false in both directions: concepts may have nonlinguistic structure, and words may have structure that is not part of the corresponding concept. These considerations suggest that much work needs to be done in clarifying the relation between words and concepts. I am not alone in

holding this view. Many researchers (some of whom I have criticized at length) are in agreement about this. Fodor, for instance, admits that 'Getting clear on the word-concept relation is no small matter' (Fodor, 1995, p. 34). In a similar vein, Robert Matthews casts doubt on whether our propositional attitudes have a structure that is mirrored in any interesting way by our natural language attitude reports (Matthews, 1994, 2002).

6. Conclusion

In the course of this paper, we have seen several reasons for holding that many of our words are structured, and that processes in the lexicon produce these structures or otherwise operate on them. One can't maintain, for instance, that such theories will never explain impossible words on the grounds that any word can enter the language as a primitive. Such a theory results in methodological impotence and empirical falsity. Similarly, one can't do away with the process of lexicalization. The process is crucial to the theories it fits into, and the attempt to replace it with a nomic connection view shows no promise. Likewise, one can't maintain that words are unstructured simply by redefining the lexicon so that everything happens in the syntax. Such a strategy relabels things without solving any of the original problems. Finally, we saw that a fundamental assumption linking theories of words to theories of concepts is badly in need of clarification and justification. This presents us with an interesting challenge, since our words surely express concepts.

In this paper, I have tried to expose some of the complexity and considerations that go into constructing a theory of words. By examining a few linguistic details, I hope to have made a small contribution to a theory of the word-concept relation. When this relation is better understood, we may be able to address some of the questions that drive psychologists and philosophers of mind. Until then, we are not in a position to make many justified claims about what word-structure teaches us about concepts.

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

References

- Baker, M. 1988: *Incorporation*. Chicago: University of Chicago Press.
Baker, M. 1996: *The Polysynthesis Parameter*. Oxford: OUP.
Baker, M. 1997: Thematic roles and syntactic structure. In L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer.
Baker, M., Johnson, K., and Roberts, I. 1989: Passive arguments raised. *Linguistic Inquiry*, 20, 219–52.

- Bittner, M and Hale, K. 1996: Ergativity: toward a theory of a heterogeneous class. *Linguistic Inquiry*, 27, 531–604.
- Bloom, P. 2000: *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.
- Carter, R. 1976: Some constraints on possible words. *Semantikos*, 1, 27–66.
- Chomsky, N. 1981: *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky N. 1986: *Knowledge of Language*. Westport, Conn.: Praeger.
- Chomsky N. 2000: Minimalist inquiries: the framework. In R. Martin *et al.* (eds), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.
- Chomsky N. 2001: Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale*. Cambridge, MA: MIT Press.
- Comrie, B. 1985: Causative verb formation and other verb-deriving morphology. In T. Shopen (ed.), *Language Typology and Syntactic Description*. Vol. III. Cambridge: CUP.
- Dowty, D. 1979: *Word Meaning and Montague Grammar*. Dordrecht: D. Reidel.
- Dowty, D. 1989: On the semantic content of notion of 'thematic role'. In G. Chierchia, B. Partee and Turner, R. (eds), *Properties, Types and Meaning*. Volume II. Dordrecht: Kluwer.
- Dowty, D. 1991: Thematic proto-roles and argument selection. *Language*, 67, 547–619.
- Fodor, J. 1970: Three reasons for not deriving 'kill' from 'cause to die'. *Linguistic Inquiry*, 1, 429–438.
- Fodor, J. 1995: Review of Christopher Peacocke's *A Study of Concepts*. *The London Review of Books*. April 20. Reprinted in J. Fodor, 1998: *In Critical Condition*. Cambridge, MA: MIT Press.
- Fodor, J. 1998: *Concepts*. Oxford: OUP.
- Fodor, J. and Lepore, E. 1999: Impossible words? *Linguistic Inquiry*, 30, 445–53.
- Fodor, J. and Lepore, E. 1998: The emptiness of the lexicon: reflections on James Pustejovsky's *The Generative Lexicon*. *Linguistic Inquiry*, 29, 269–288.
- Fodor, J. and Lepore, E. 1992: *Holism: A Shopper's Guide*. Oxford: Blackwell.
- Gleitman, L. 1990: The structural sources of verb meanings. *Language Acquisition*, 1, 3–55.
- Grimshaw, J. 1990: *Argument Structure*. Cambridge: MIT Press.
- Grimshaw, J. 1997: Locality and extended projection. In P. Coopmans, M. Everaert and J. Grimshaw (eds), *Lexical Specification and Insertion*. Hillsdale, NJ: Lawrence Erlbaum.
- Gropen, J., Pinker, S., Hollander, M. and Goldberg, R. 1991: Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, 18, 115–151.
- Gropen, J., Pinker, S., Hollander, M. and Goldberg, R. 1991a: Affectedness and direct objects: the role of lexical semantics in the acquisition of verb argument structure. *Cognition*, 41, 153–195.
- Hale, K. and Keyser, S. J. 2003: *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.

- Hale, K. and Keyser, S. 1999: A response to Fodor and Lepore, 'Impossible words?'. *Linguistic Inquiry*, 30, 453–66.
- Hale, K. and Keyser, S. 1997: On the complex nature of simple predicators. In A. Alsina, J. Bresnan, and P. Sells (eds), *Complex Predicates*. Stanford: CSLI Publications.
- Hale, K. and Keyser, S. 1993: On argument structure and the lexical expression of syntactic relations. In K. Hale and S. Keyser (eds), *The View from Building 20*. Cambridge, MA: MIT Press.
- Hale, K. and Keyser, S. 1987: A view from the middle. *Lexicon Project Working Papers*, 10, Cambridge, MA: Center for Cognitive Science, MIT.
- Halle, M., and Marantz, A. 1993: Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser (eds), *The View from Building 20*. Cambridge, MA: MIT Press.
- Jackendoff, R. 1990: *Semantic Structures*. Cambridge, MA: MIT Press.
- Kornfilt, J., and Correa, N. 1993: From conceptual structure to syntax: projecting from resultatives. In E. Reuland and W. Abraham (eds), *Knowledge and Language, Vol. II: Lexical and Conceptual Structure*. Dordrecht: Kluwer.
- Kratzer, A. 1996: Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds), *Phrase Structure and the Lexicon*. Dordrecht: Kluwer.
- Levin, B. 1993: *English Verb Classes and Alternations*. Chicago: University of Chicago Press.
- Marantz, A. 1997: No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics*, 4.2, 201–225.
- Margolis, E. and Laurence, S. 1999: Concepts and cognitive science. In Laurence and Margolis (eds), *Concepts: Core Readings*. Cambridge, MA: MIT Press.
- Matthews, R. 1994: The measure of mind. *Mind*, 103, 131–146.
- Matthews, R. 2002: Logical form and the relational conception of belief. In G. Preyer and G. Peter (eds), *Logical Form and Language*. Oxford: Clarendon Press.
- McCawley, J. 1968: Lexical insertion in a transformational grammar without deep structure. In B. Darden, C. Bailey and A. Davison (eds), *Fourth Regional Meeting of the Chicago Linguistics Society*. Chicago: University of Chicago Press.
- Medin, D. 1989: Concepts and conceptual structure. *American Psychologist*, 44, 1469–1481.
- Miller, G. 1996: *The Science of Words*. New York: Scientific American Library.
- Peacocke, C. 1992: *A Study of Concepts*. Cambridge, MA: MIT Press.
- Pietroski, P. 1998: Actions, adjuncts, and agency. *Mind*, 107, 73–111.
- Pietroski, P. 2000: *Causing Actions*. Oxford: OUP.
- Pinker, S. 1989: *Learnability and Cognition*. Cambridge, MA: MIT Press.
- Pinker, S. 1999: *Words and Rules*. New York: Basic Books.
- Smith, E. 1995: Concepts and categorization. In E. Smith and D. Osherson (eds), *Thinking: An Invitation to Cognitive Science*. Cambridge, MA: MIT Press.
- Tanenhaus, M., Boland, J., Mauner, G. and Carlson, G. 1993: More on combinatory lexical information: thematic structure in parsing and interpretation. In G. Altmann and R. Shillcock (eds), *Cognitive Models of Speech Processing*. Hillsdale, NJ: Lawrence Erlbaum.

- Travis, L. 1984: *Parameters and Effects of Word Order Variation*. Ph.D. diss. MIT: Cambridge.
- Travis, L. 2000: Event structure in syntax. In C. Tenny and J. Pustejovsky (eds), *Events as Grammatical Objects*. Stanford: CSLI Publications.
- Zubizarreta, M. 1987: *Levels of Representation in the Lexicon and in the Syntax*. Dordrecht: Foris.