

Impossible Words?

Jerry Fodor and Ernest Lepore

1 Introduction

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," and it unites the generative and interpretive traditions of grammatical analysis. Hale and Keyser (HK) (1993) have endorsed a version of lexical decomposition according to which "denominal" verbs are typically derived from phrases containing the corresponding nouns: 'sing_{vtr}' is supposed to come from something like DO A SONG; 'saddle_{vtr}' is supposed to come from something like PUT A SADDLE ON; 'shelve_{vtr}' is supposed to come from something like PUT ON A SHELF; and so forth.¹ Their case for these claims revives a form of argument, the "impossible-word" argument, which has for some time been in eclipse. In this article we will claim that, whatever the right account of lexical decomposition eventually turns out to be, impossible-word arguments are infirm and should be discounted. We will use HK's paper as a text for this critique.

HK's arguments for the derivational analysis of denominal verbs are all variations on the same general theme: the derivational account predicts/explains the intuitive impossibility of certain verbs by showing that the syntactic processes that would be required to derive them are prohibited by independently well confirmed grammatical constraints. We think there are principled objections to this form of argument. So, rather than dispute over details, we will take HK's word about what grammatical constraints are independently well confirmed; if they say that a derivation is blocked, we will assume that indeed it is. (We remark in passing, however, that a number of our colleagues have cautioned us against conceding (e.g.) that structures like (9) (see page XXX below) violate the Empty Category Principle (ECP).)

Since we think that the logic of impossible-word arguments—hence their limitations—has not been made clear in the literature, we start by trying to be explicit about how such arguments are supposed to run. Three points about them will be assumed throughout our discussion.

¹ Cited forms are in single quotation marks and expressions in italics either stand for canonical names of meanings (or are emphasized expressions). We use capital letters for canonical names of semantic representations. (So, according to the decompositional view of causatives, 'kill' derives from CAUSE TO DIE and means *cause to die*.) However, for convenience, we will sometimes not distinguish between a semantic representation and the corresponding surface phrase; for example, we will say that 'kill' is derived from 'cause to die' to abbreviate the view that 'kill' and 'cause to die' are both derived from CAUSE TO DIE. Corner quotes ($\lceil \cdot \rceil$) enclose variables that range over italicized formulas.

First, the *conclusion* of an impossible-word argument is that there is plausible evidence for a grammatical rule/operation of *lexicalization*. (For exposition, we assume that lexicalization is a transformation.) The defining property of lexicalization is that it derives a surface lexical item from an underlying semantic representation. In the cases of interest, (a) the presumed underlying representation has internal constituent structure (in effect, it is a phrase), and (b) the surface item is more or less synonymous with the semantic representation from which it is derived.²

Second, the basic form of argument for lexicalization is that it satisfies independently motivated constraints on syntactic rules. Since it does, a grammar that fails to recognize lexicalization thereby misses a certain generalization, namely, that the formation rules for lexical items are a species of the formation rules for overtly complex linguistic structures like phrases.

Third, the evidence for the claim that lexicalization satisfies independently motivated grammatical constraints is that lexicalizations that violate them produce intuitively "impossible words." Impossible words correspond not to *de facto* gaps in the lexicon but to expressions that could not occur in a language *modulo* the grammatical universals.³ In fact, all of HK's examples are of the following form (see pp. 60-64): they present a sentence that contains a neologistic verb, together with a paraphrase; and what is explained by appeal to the putative independently motivated grammatical constraints is why the former could not be derived from the latter (more precisely, why the former could not be derived from the representation that directly underlies the latter). HK say that the existence of constraints that block these derivations explains why " . . . English simply does not have verbs . . . that . . . have meanings corresponding more or less to the . . . paraphrases given here" (pp. 59-60).

We especially wish to emphasize how this sort of argument fits into the discussion about lexical decomposition (i.e., about whether surface lexical items sometimes have complex semantic representations underlying them). From that perspective, the question of interest is not whether there are impossible words; it is not even how to explain why impossible words are impossible. Instead, it is this: *Does the putative explanation of why they are impossible offer an argument for lexicalization?* That is, does it offer an argument that some surface lexical items are derived from underlying phrasal representations? We will argue that it doesn't and couldn't.

² There is considerable unclarity in the lexical semantics literature about just what meaning relations are required to hold between a lexical item and the semantic representation from which it is derived. Some claim, for example, that such derivations capture only the core or paradigmatic meaning(s) of surface lexical items. We will be neutral on this issue. We are concerned with the evidential status of the claim that surface lexical items *ever* derive from underlying phrases, whatever semantic relations such derivations are supposed to preserve.

³ We are assuming that the constraints in question are universal. A weaker view might relativize the impossibility of lexical items to particular languages (e.g., *w* is impossible *in English* if its derivation would require rules or structures that English grammar happens not to acknowledge). As we understand them, impossible-word arguments generally have the stronger notion of constraint in mind; but the considerations we will raise do not depend on the distinction.

Prima facie, there are two kinds of ways in which constraints on possible lexical items might be predicted from general properties of grammatical derivations. Roughly, it might turn out that the derivation of a lexical item is blocked because it would violate some independently motivated constraint on the *kinds of grammatical rules* there can be, or because it violates some independently motivated constraint on the *kinds of structures* that grammatical rules are allowed to generate.⁴ (Though they do not discuss this difference, HK offer examples of both kinds.) It will be convenient for our exposition to consider these two kinds of blocked derivations separately since there are slightly different reasons why neither can provide evidence for lexicalization. But it's well to keep the bottom line in mind: we claim that neither of these kinds of impossible-word argument can work. And since, as far as we can tell, these two are all the kinds of impossible-word arguments there could be, we also claim that impossible-word arguments are infirm in principle: *no impossible-word argument could be evidence for lexicalization*.

2 Impossible Rules

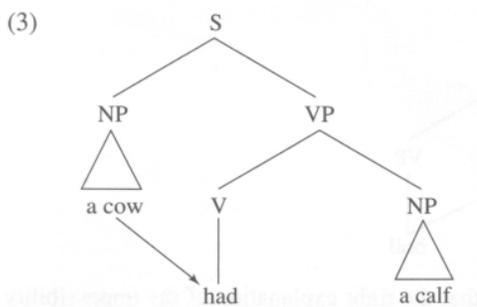
In case 1 the derivation is blocked because a rule (a movement rule in all the examples that HK give) that the derivation depends on is impossible. For example, HK claim that there could not be a verb 'to cow' such that (1) is well formed and means *a cow had a calf*.

(1) *It cowed a calf.

They explain this as follows. Since, by assumption, lexical items typically derive from the structures that underlie their phrasal synonyms, sentence (1) would presumably come from the same structure that underlies sentence (2).

(2) A cow had a calf.

The structure of (2) is presumably (3), and the derivation would crucially depend on a transformation of subject lowering (followed by insertion of dummy *it* in the surface subject position).



But, as HK remark, "[i]t is well known that a subject. . . cannot incorporate into the verb that heads its predicate" (p. 60). Since subject lowering is an illicit

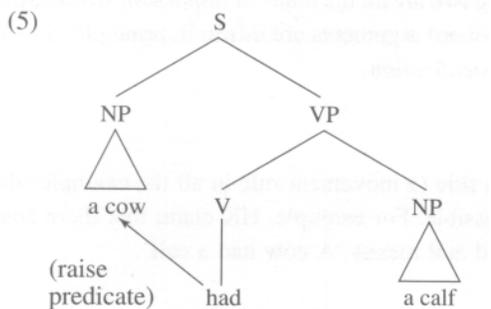
⁴ This corresponds to the familiar distinction between derivations that block because they do not terminate and derivations that block because they generate structures that are subsequently filtered.

operation, the putative derivation (3) is impossible; hence the intuition that (1) is likewise impossible.

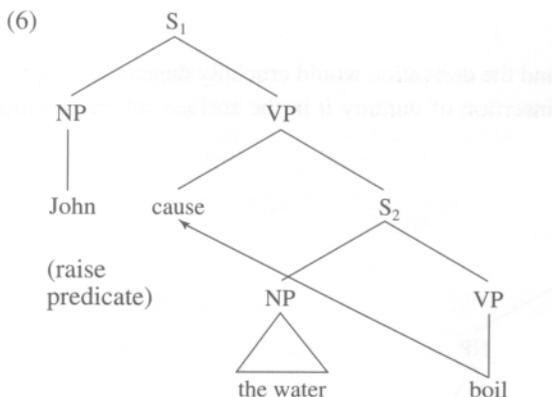
A preliminary comment: We think that this account is not plausible even independent of the general issues about lexical decomposition. In effect, HK trace the nonexistence of (1) to a (presumably universal) prohibition against transformations that lower subjects into predicates. Notice, however, that (4) is also impossible on the reading (1).

(4) *A cow did a calf.

However, the derivation of (4) requires, not lowering the subject NP, but raising the predicate (followed by insertion of dummy 'do' in the surface verb position). Since (4) is impossible, the derivation (5) presumably has to be blocked.



But presumably (4) cannot be blocked by a prohibition against raising predicates, since such raising is required to derive (e.g.) causative transitives from their intransitive counterparts, as in (6). (6) is, in fact, a kind of derivation of which HK approve (see p. 79).



The symmetry between (1) and (4) suggests that the right explanation of the impossibility of the former is not that subject lowering is illegal. Rather, it is that deriving either (1) or (3) from (4) would involve lexicalizing an expression that itself is not a well-formed constituent, namely, [cow did]. That lexicalizing a nonconstituent is illegal would presumably follow just from the assumption that lexicalization is a transformation; in any case, it is widely taken for granted in the generative semantics literature that only constituents can be lexicalized.

But we do not insist on this analysis. We are about to present an argument that neither examples where derivations block because the rules that they employ are illegal nor derivations that block because the structures that they generate are illegal can provide evidence for lexicalization. Here is our argument against the first disjunct.

There must be something wrong with HK's account of cases like (1) since, even if it did explain why there couldn't be a *derived verb* 'to cow' with the paraphrase (2), it does not explain why there couldn't be a *primitive*, underived verb 'to cow' with the paraphrase (2). As far as we can tell, this sort of point applies to any attempt to explain why a word is impossible by reference to the impossibility of a certain transformation. 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 $\lceil e \rceil$ 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 primitives.⁵ We assume, along with HK, that the intuition about (1) is that it is *impossible*—and not just that *if* it is possible, then it is underived. (We do not suppose that anyone, except perhaps linguists, has intuitions of the latter kind.) So we claim that HK have not explained the intuition that 'to cow' is impossible.

We repeat that this sort of objection holds in any case where a derivation is said to be blocked by the lack of transformations that would construct an appropriate domain for lexicalization—as, for example, the lack of a subject-lowering transformation is supposed to explain why there is no 'the cow had' to provide a domain for the introduction of 'cowed'. Assuming that *a cow had* is a well-formed meaning, HK's sort of account does not explain why there could not be a primitive lexical item 'cowed' that means *a cow had*. Likewise in the general case: assuming that 'e' is a well-formed meaning, the assumption that the derivation of the corresponding expression $\lceil e \rceil$ is blocked does not explain why 'e' cannot be the meaning of a primitive lexical item, whether or not $\lceil e \rceil$ is syntactically derivable.

It may be well (in passing) to explicitly distinguish the two kinds of theories that we mentioned in [footnote 2](#). Presumably $\lceil e \rceil$ is underivable either because the grammar of English lacks the relevant rules or because no grammar of a natural language could have the relevant rules. In the first case, the most that would follow is that, if w is a lexical primitive, then it cannot be defined *in English* (i.e., there is no phrase in English that is synonymous with w). In the second case, the most that would follow is that if w is a lexical primitive, then it cannot be defined *in any natural language*. Further argument would then be required to show that if the meaning of w cannot be defined in any natural language, w is *ipso facto* an impossible word. (It is not obvious that such arguments would be forthcoming given the empirical plausibility of the claim

⁵ We put it this way because, strictly speaking, a demonstration that $\lceil e \rceil$ is not well formed could only show that if w is possible and derived, it cannot be derived from $\lceil e \rceil$. Notice that this would leave open not just that w is derived, but that it is derived *and means* 'e': the derivational source of w might be some phrasal synonym of $\lceil e \rceil$.

that the majority of lexically primitive expressions in every natural language are not definable.)

It is important to bear in mind, at this point, what it seems to us that impossible-word arguments are forever forgetting: 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.) But how could any conclusion of that sort follow from a demonstration that ‘*e*’ isn’t derivable, from the grammar of a (or any) natural language? Maybe it’s derivable in Fortran. We suspect that impossible-word arguments suffer from a use/mention confusion—in particular, confusing “the expression ‘*e*’ isn’t possible” (which is what instances of the argument actually show when they are successful) with “the meaning ‘*e*’ isn’t possible” (which is what using impossible-word arguments to defend lexicalization would actually require). There may be sound arguments the conclusion of which has the form “‘*e*’ is not a possible meaning” but we don’t know what they would look like; nor have we ever seen one. (Notice that even ‘round square’ is a possible meaning; in fact, it’s what *round square* means in English.)

3 Impossible Structures

In case 2 the derivation is blocked because the structure it generates is impossible. (The constraint is on the product rather than the process.)

HK do not always assume that what blocks derivations in their examples is a prohibition against moving a certain constituent to (or from) a certain position (e.g., the putative principle that there cannot be a transformation that lowers a subject into a predicate). Instead, sometimes the constraint is on the form of the tree that a derivation *would* produce if it were legal. Here is HK’s explanation of why there cannot be a verb ‘shelve . . . on’ such that (7) is grammatical and means (8).⁶

(7) *He shelved the books on.

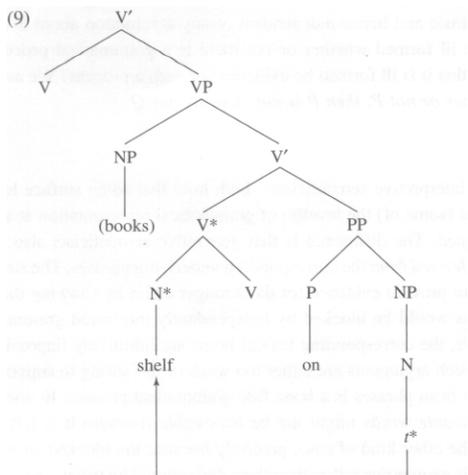
(8) He put the books on a shelf.

Each of these hypothetical items, ‘shelve (books) on . . .’, is derived by incorporation of the noun that heads the complement of the preposition, as shown in [(9)]. The trace of incorporation is thus “too far” from its antecedent and is therefore not properly governed, violating the ECP.

Although the trace is coindexed with the verb to which its antecedent is

⁶ Notice that (8) expresses a possible meaning (in fact, it expresses the meaning that ‘He put the books on the shelf’ actually has). This seems to be a clear case where the question “What expressions are syntactically derivable?” has been confused with the question “What lexical items are (semantically) possible?” However, the discussion to *follow* does not rely on this criticism. (It adds to the confusion that expressions like ‘phone up’ may well be complex words, that is, lexical items with internal constituent structure. This *is* not, however, the sort of analysis HK are imagining for ‘shelve . . . on’ in (7). In particular, the structure HK are considering *is* one where ‘the shelf’ *is* the object *of* the preposition ‘on’, not one where it *is* the object *of* a complex verb ‘shelve on’.)

adjoined (as indicated by the asterisk notation), this verb does not govern the trace. The preposition is a "closer governor," defining PP as the minimal governing domain for the trace By Minimality, therefore, PP is a barrier to government from the more distant verb. (p. 61)



In short, (8) cannot come from (7) because (7) is illegal; and (7) is illegal because (9) violates a constraint on variable binding that well-formed formulas are *ipso facto* required to comply with. Notice, crucially, that this explanation would account for why (7) is ungrammatical whether or not the putative verb ‘shelve . . . on’ is supposed to be derived—a *fortiori*, whether or not (7) is supposed to be derived from (8). In particular, the ECP would block (7) even if ‘shelve . . . on’ is primitive and is base-generated. So, in this sort of case, HK's explanation of the "impossible" form really does explain why it is *impossible*. Compare case 1, where the fact that the derivation of (1) blocks shows only that if ‘to cow’ is possible, then it must be primitive. So HK's account of case 2 has the virtue that their account of case 1 crucially lacks.

However, it also has the defects of this virtue. Remember that the whole point of making an impossible-word argument is to provide evidence for lexical lexicalization—that is, to provide evidence for inferring that there is a grammatical process that derives surface lexical items from complex, phrasal underlying structures. So, 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 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 how would this pattern of inference go in the present case? If (7) is ill formed for the reason that HK allege—namely, because there is a mandatory constraint that the structure (9) fails to meet—then, patently, (7) is ill formed whether or not there is a lexicalization transformation. Indeed, on the present assumption, the ill-formedness of (7) depends entirely on features of its geometry; it is, as it were, intrinsic and hence independent of any assumption about how (7) was derived. But if (7) 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? We assume it is not in dispute that if Q *whether or not P*, then *P is not evidence for Q*.

4 Conclusion

"Generative semanticists" and "interpretive semanticists" both hold that some surface lexical items are represented by phrases at (some of) the level(s) of grammatical representation at which semantic interpretations are assigned. The difference is that generative semanticists also think that such surface lexical items are *derived from* the corresponding underlying phrases. The strategy of impossible-word arguments is to provide evidence for the stronger claim by showing that the derivations of certain lexical items would be blocked by independently motivated grammatical constraints and that, when they are, the corresponding lexical items are intuitively impossible.

We claim, however, that all such arguments are either too weak or too strong to support the thesis that the derivation of words from phrases is a bona fide grammatical process. In one kind of case, it is left open that *underivable* words might not be *impossible* (because it is left open that they might be primitive); in the other kind of case, precisely because the blocked structures are ungrammatical in virtue of their geometries rather than their derivational histories, their being ill formed provides no evidence for or against there being such a process as lexicalization.

But quite aside from the relatively mechanical questions that have primarily concerned us, there is a prima facie reason to think that impossible-word arguments must be deeply flawed. In effect, such arguments infer conclusions of the form "'e' is not a possible word meaning" from premises of the form "[e] is not derivable in a certain grammar." Now it is true that, once upon a time, philosophers dreamt of a "logically perfect" language in which all and only the *meaningful* (*verifiable?*) (*truth evaluable?*) sentences would be *syntactically* well formed (see, e.g., Wittgenstein in *The Tractatus*). But even philosophers never thought that one could formalize English in such a language; the most they hoped for was to formalize mathematics or physics. In any case, these days everybody thinks that both projects are hopeless. Our point is that the inference from "[e] is underivable" to "'e' is not a possible meaning" would be sound only on the assumption that [e] is an expression in such a language. So cognitive scientists who endorse impossible-word arguments are thereby committed to a kind of linguistic utopianism that even philosophers are unable to believe.

Barring blatant confusions of use with mention, such facts as that there is no rule of subject raising, or that the ECP controls variable binding, say nothing at all about what words are possible. All they tell us is exactly what they seem to—namely, why sentences like (1) and (7) are not grammatical.

Reference

Hale, Kenneth, and Samuel Jay Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In *The View from Building 20: Essays on Linguistics in Honor of Sylvain Bromberger*, ed. Kenneth Hale and Samuel Jay Keyser, 53-109. Cambridge, Mass.: MIT Press.