evidence for lexicalism:

a critical review

michael a. covington

IULC

EVIDENCE FOR LEXICALISM:

A CRITICAL REVIEW

by

Michael A. Covington

Yale University

TABLE OF CONTENTS

0. Introduction

The Lexicalist Hypothesis is, roughly, the hypothesis that both members of pairs of words such as *decide:decision* and *tie:untie* are represented in the lexicon (and shown in some way to be related), rather than that there is a transformation converting one into the other; in other words, "transformations do not perform derivational morphology" (Jackendoff 1972:12-13).

The major arguments that have been put forward in favor of this hypothesis fall into two classes: arguments that the Lexicalist Hypothesis in itself is preferable to other accounts of derivation (Chomsky 1970; for counterarguments see Ross 1973, Postal 1974, and McCawley 1975), and arguments that on the foundation of the Lexicalist Hypothesis it is possible to construct a theory of derivation that expresses significant generalizations not otherwise easily captured (Jackendoff 1975, Aronoff 1976). In this paper I shall review and critique these arguments, endeavoring to place each argument in context and adding data and argumentation of my own whenever appropriate.[1]

Chomsky proposed the Lexicalist Hypothesis in the context of a grammatical theory that includes syntactic features, a semantic component that determines grammatical relations exclusively from deep structure, no unrecoverable deletions, and no prelexical transformations. Throughout, I shall continue to assume such a theory. Further, I shall confine myself to arguments for and against a lexicalist treatment of derivation; the proposals of Freidin (1975), Bresnan (1978), and others to account for phenomena such as Passive and Dative Movement by means of lexical rules are outside my central area of concern, though relevant to it.

1. The Lexicalist Hypothesis per se

1.1 Chomsky's original arguments

In his 1970 paper "Remarks on Nominalization", Chomsky formulates the Lexicalist Hypothesis as follows:

> Let us propose then, as a tentative hypothesis, that a great many items appear in the lexicon with fixed selectional and strict subcategorization features, but with a choice as to the features associated with the lexical categories noun, verb, adjective. The lexical entry may specify that semantic features are in part dependent on the choice of one or another of these categorial features. (Chomsky 1970, p. 190)

Since Chomsky does not elaborate this branching-lexical-entry mechanism, other lexicalists (Jackendoff 1975, for instance) have felt free to replace it with another model in which derivationally related forms have completely separate lexical entries, related by rule.

In either case, the Lexicalist Hypothesis, per se, has three major consequences:

(a) Derivation precedes all transformations. The output of a transformation can never be the input to a derivational process, since derivational processes take place in the lexicon.

(b) A derived form is the same part of speech in deep structure and on the surface; there is no stage of derivation at which, for instance, *decision* is a verb.

(c) Since derivationally related forms are distinct in the lexicon, and since the lexicon is in general the repository of unpredictable properties of lexical items, it follows that, in the absence of a constraint to the contrary, derivationally related forms can be expected to differ from each other in unpredictable ways (though they need not do so).

Corresponding to these three consequences, Chomsky draws three contrasts between derived nominals (*John's proof of the theorem*, *Max's criticism of the book* and gerundive nominals (*John's proving the theorem*, *Max's criticizing the book*). For expository convenience I shall discuss these three contrasts, and the arguments based on them, in a different order than Chomsky. The contrasts are:
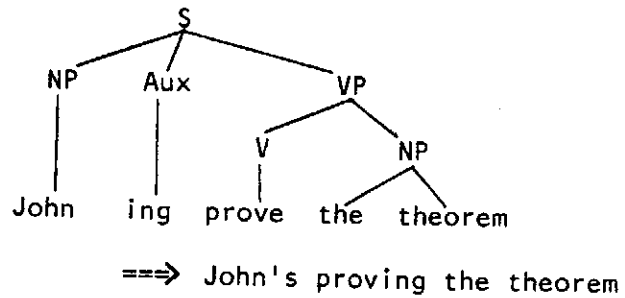
(1) There exist gerundive nominals corresponding to a full range of sentences, including sentences that have undergone Raising, Dative Movement, and the like, but derived nominals in general correspond only to base sentences.

(2) Gerundive nominals occur in structures similar to those of sentences, while derived nominals occur in phrases with the structure of non-sentential NPs.

(3) The meanings and morphological properties of gerundive nominals are completely predictable from the verbs from which they are derived. Derived nominals, on the other hand, often have idiosyncratic properties; for instance, *creation* pluralizes but *destruction* does not, and *edition* can refer to the product of a process that is not, strictly speaking, describable as *editing* (cf. *hardback and paperback editions* that differ only in their covers).

To account for these contrasts, Chomsky proposes that gerundives are derived transformationally from embedded sentences that have a special gerundive marker in place of the tense, while derived nominals are listed separately in the lexicon and inserted into (non-sentential) NP structures in the base. His proposed deep structures for *John's proving the theorem* and *John's proof of the theorem* are something like the following (see Jackendoff 1977, p. 10-1):

(1.1-1)

```
                          S
            _____/_____
          NP      Aux          VP
          |        |         /\
          |        |        V    NP
          |        |        |   /\
        John     ing     prove the  theorem
```

==⟩ John's proving the theorem

(1.1-2)

```
                      NP
              ____/|_____
           Det      N          PP
            |       |        /\
          Poss      |       P    NP
          /\        |       |   /\
        NP   \      |       |  /\
        |     \     |       | the theorem
      John    's  proof    of
```
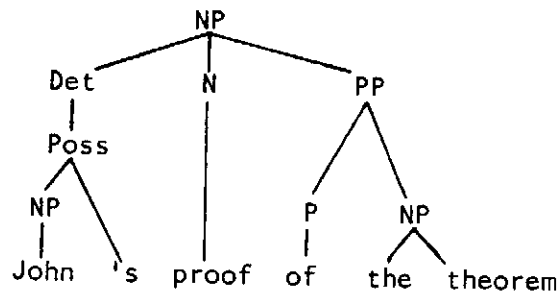
Assuming that Chomsky's observations are correct, these analyses explain the situation neatly: derived nominals ARE nouns, with their own lexical entries, while gerundives are derived from sentences. But the burden of showing that (for example) *proof* and *prove* are related is now thrown onto the lexicon, and, if one accepts the Lexicalist Hypothesis, a new and potentially powerful device has to be added to the grammar. The question is whether this is a well-motivated move.

Let us now examine Chomsky's arguments in detail.

1.1.1 Syntactic restrictedness
Corresponding to sentences that have undergone various syntactic transformations, there are gerundive nominals but no derived nominals. For instance:

(1.1-3) TOUGH-movement:

    (a) John is easy to please. (cf.: It is easy to please John.)

    (b) John's being easy to please

    (c) *John's easiness to please (cf.: the easiness of pleasing John)

(1.1-4) Subject Raising

  (a) John is certain to win. (cf.: It is certain that John will

  (b) John's being certain to win   win.)

  (c) *John's certainty to win (cf.: the certainty that John
                                           will win)[2]

(1.1-5) Dative Movement

  (a) John gave Bill a book. (cf.: John gave a book to Bill.)

  (b) John's giving Bill a book

  (c) *John's gift of Bill a book (cf.: John's gift of a book
                                             to Bill)

(1.1-6) Particle Movement

  (a) John looked the information up. (cf.: John looked up the

  (b) John's looking the information up   information.)

  (c) *John's looking of the information up

                        (cf.: John's looking up of the
                                 information)

This is just what Chomsky's analysis predicts. Gerundives start out as
embedded sentences in deep structure and thus can undergo transformations
that apply to sentences; but since nominalization is accounted for by the
lexicon rather than the transformational component, derived nominals are
not sentences at any level of syntactic representation. This is an
instance of the more general prediction that derivational processes never
follow transformations. But Newmeyer (1971) points out that such a
prediction follows not only from the Lexicalist Hypothesis, but also from
the much less radical hypothesis that THE TRANSFORMATIONS THAT PERFORM
DERIVATION ARE ORDERED BEFORE ALL OTHER TRANSFORMATIONS. If a precyclic
rule of Nominalization converts embedded sentences into non-sentential
NPs, then the data are accounted for without recourse to lexicalism.

    For that matter, does the generalization that derived nominals do
not undergo sentential transformations in fact hold up? Chomsky (1970:202ff.)
deals with one apparent counterexample himself: the case of passive nominals
such as *the city's destruction by the enemy*. He argues convincingly that
there is independent evidence for one or more transformations that 'passivize'
NPs, relating the examples in the sets below. (A strong case can be made
for taking the (b) structures rather than the (a) structures as basic, but
this does not affect out present argument; note that it has likewise been
proposed that passive sentences have the *by*-phrase present in deep structure.)

(1.1-7) (a) the enemy's destruction of the city

(b) the destruction of the city by the enemy

(c) the city's destruction by the enemy

(1.1-8) (a) Holbein's portrait of Henry VIII

(b) the portrait of Henry VIII by Holbein

(c) Henry VIII's portrait by Holbein

Suppose that (1.1-7 a, b, c) are derived from active and passive versions of a deep structure containing *destroy*. Then, to capture the generalization that (1.1-7) and (1.1-8) are parallel, it is necessary to posit that the latter set comes from a deep structure containing an abstract verb that nominalizes as *portrait*. Such a treatment becomes implausible in view of the variety of abstract verbs that would be needed to cover 'active'-'passive' pairs like these:

(1.1-9) (a) Frank Lloyd Wright's houses:

houses by Frank Lloyd Wright

(b) Tolkien's book about hobbits:

the book about hobbits by Tolkien

(c) Calvin Klein's blue jeans:

blue jeans by Calvin Klein

Do we really want an abstract verb that nominalizes as *blue jeans*? It seems much more reasonable to propose that certain classes of English nouns, including words for artistic creations and also most derived nominals, can take agent phrases within the NP, and that the sets of phrases above are related by one or more transformations. Once such transformations are available for (1.1-8) and (1.1-9), they can also be used to relate 'active' and 'passive' derived nominals, as in (1.1-7), and it is not necessary to invoke sentential Passive.[3] (Whether this newly proposed process of NP-Passive will turn out to be generalizable with sentential Passive is then a moot point. Chomsky (1970) expects that it will; McCawley (1975:215-6) argues that it will not.)

Ross (1973) points out that not only passivization, but also Equi and a rule he calls PP Shift, seem to apply to NPs as well as sentences (assuming the Lexicalist Hypothesis). Corresponding to the fairly clear case of Equi in (1.1-10), we have the nominalization in (1.1-11):

(1.1-10) Thom is ready for [for Thom to operate on Sue]

===> Thom is ready to operate on Sue.

(1.1-11)   Thom's readiness for [for Thom to operate on Sue]

==➤   Thom's readiness to operate on Sue

(There has to be a later rule that reduces the two *fors* to one.) That Equi is needed in NPs that are not nominalizations is suggested by phrases like *Nan's yen to go around the world*. The following are examples of PP Shift in a sentence and in a noun phrase:

(1.1-12)   (a)   We talked with Gretchen about hockey.

We talked about hockey with Gretchen.

(b)   the message from Ghent to Aix

the message to Aix from Ghent

Yet Dative Movement, which superficially resembles PP Shift, does not apply in NPs (1.1-5, above); and Object Deletion, which bears a great resemblance to Equi, does not apply in NPs, as shown in the examples below:

(1.1-13)   (a)   Thom is ready for [for Sue to operate on Thom]

==➤   Thom is ready for Sue to operate on.

(b)   Thom's readiness for [for Sue to operate on Thom]

==➤   *Thom's readiness for Sue to operate on

Ross's point is that NPs are not exempt from ALL sentential transformations, only SOME of them; and in the absence of a principled way to decide which transformations belong in which group, lexicalism provides no real explanation of the phenomena. All that is left is the much less striking generalization (quoting from Ross 1973:214):

(1.1-14)   The rules which apply to NPs whose head noun is
morphologically complex will not differ from
those that apply within NPs whose head noun is
morphologically simple.

From this Ross somehow goes on to conclude that the whole argument is "without force in choosing between the two theories" (i.e., lexicalism and the transformational account of nominalization). This simply does not follow. Lexicalism predicts that derived nominals will behave exactly like NPs throughout the syntactic derivation, whereas the transformational account predicts that they will behave like sentences before Nominalization and like NPs afterward. The import of (1.1-14) is that morphologically complex nouns (e.g., derived nominals) always act like simple nouns, not like sentences. So if (1.1-14) is upheld, the predictions of lexicalism are borne out.

Postal (1974:322-55) states a much more forceful objection to lexicalism, which according to Newmeyer (1980:119) has been known since the late 1960s. In phrases like

(1.1-15)   John's tendency to sleep late

Murphy's continuation as police chief

his estimate of your weight as (being) 200 pounds

Nominalization appears to be fed by Raising, which is impossible if Nominalization is not a transformation.

For the sake of the argument I shall assume that Postal is right in proposing Raising analyses of *tend*, *continue as*, and *estimate* in the relevant constructions, though this is obviously open to question; the parallelism of the syntax of these verbs with that of the nominalizations derived from them (as in the examples above) then suggests that the nominalizations likewise involve Raising. Postal's analyses are, roughly, the following (using square brackets to enclose embedded clauses):

(1.1-16)   Raising in sentences:

(a)   tend [for John to sleep late]

==>   John tends [to sleep late]

(b)   continue [Murphy be police chief]

==>   Murphy continues [as police chief]

(c)   he estimated [that your weight was 200 pounds]

==>   He estimated your weight [as (being) 200 pounds].

(1.1-17)   The corresponding nominalizations:

(a)   John's tendency [to sleep late]

(b)   Murphy's continuation [as police chief]

(c)   his estimate of your weight [as (being) 200 pounds]

As Postal notes, the strongest syntactic test for Raising is unavailable in the case of nominalizations. In general, sentences like *there tend to be riots on campuses, it tends to rain every afternoon,* and *advantage tends to be taken of such opportunities* provide a strong argument for Raising because generalizations are lost if *there, it,* or *advantage* (in the relevant, idiomatic sense) is generated as the deep subject of *tend*. The fact that the corresponding nominalizations

(1.1-18)   *there's tendency to be riots on campuses

*its tendency to rain every afternoon

*advantage's tendency to be taken of such opportunities

do not occur is explained by the independent fact that expletive *there* and *it*, and other semantically empty NPs (such as idiom chunks), do not form possessives (we likewise do not get *I'm tired of advantage's being taken of me*). Thus, this kind of data provides no argument for or against Raising in nominalizations.

However, there are syntactic arguments that (for example) *Kronzheim* is a subject in *the tendency for Kronzheim to make all the decisions* (which has not undergone the putative Raising) but not in *the tendency of Kronzheim to make all the decisions* (which has), exactly as the Raising analysis predicts. One such argument (from Postal 1974:333,348) is based on the generalization that only subject Nps can end in *alone*. We observe the pattern:

(1.1-19)   (a)   the tendency for Kronzheim alone to make all the decisions

*the tendency of Kronzheim alone to make all the decisions

(b)   his estimate that Bob's weight alone was 200 pounds

*his estimate of Bob's weight alone as 200 pounds

(In (b), *alone* has to be taken as meaning 'only', not 'by himself'.) Such arguments rule out a conceivable analysis in which the difference between *for* and *of* in (a) is treated as purely superficial (and *Kronzheim's tendency* comes from *the tendency of Kronzheim* instead of the other way around), but they do not discriminate between Equi and Raising.

For that Postal appeals to a semantic argument.[4] Consider the examples:

(1.1-20)   (a) (Raising:)   The queen tends to be praised by the servants.

vs.   The servants tend to praise the queen.

(b) (Equi:)   The queen wishes to be praised by the servants.

vs.   The servants wish to praise the queen.

In (b) there is a clear truth-value distinction between the two sentences; it is easy to imagine a situation in which one of the sentences is true

and the other is false. In (a), on the other hand, different as the sentences may be in focus, emphasis, and even semantic or syntactic acceptability, they show no real truth-value contrast. This pattern holds up for Raising and Equi verbs in general; moreover, it also holds for the corresponding nominalizations:

(1.1-21)　(a) (Raising:)　the queen's tendency to be praised by
　　　　　　　　　　　　the servants

　　　　　　vs.　　the servants' tendency to praise the queen

　　　　(b) (Equi:)　　the queen's wish to be praised by the
　　　　　　　　　　　　servants

　　　　　　vs.　　the servants' wish to praise the queen

The phrases in (a) refer to one tendency; those in (b), to two distinct wishes, one of which can exist without the other.

In response to Postal's arguments two things must be said. First, even assuming that his analysis is correct, does Raising apply before or after Nominalization? If before, then lexicalism is indeed done for; if after, then Raising merely joins the list of transformations that (like Passive and Equi) apply to NPs as well as to sentences.

A good case can in fact be made for doing Raising after Nominalization. In applying Passive and Equi to NPs, Chomsky and Ross implicitly assumed that the possessive that precedes the head noun of an NP corresponds to the subject of a sentence, and the post-nominal of-phrase corresponds to the object. Making this same assumption, we find that Raising is exactly what is needed to derive the (b) examples below from the (a) examples or something very like them:

(1.1-22)　(a)　the tendency for John to sleep late

　　　　　　the continuation of Murphy's being police chief (?)

　　　　　　his estimate that your weight is 200 pounds

　　　　(b)　John's tendency to sleep late

　　　　　　Murphy's continuation as police chief

　　　　　　his estimate of your weight as 200 pounds

It may well be that the sources of the (b) examples have the embedded clauses marked with different complementizers than appear in the (a) examples; but there seems to be no obvious problem with allowing Raising to apply to the nominalizations rather than the sentences that (according to Postal but not Chomsky) underlie them.

In fact, with appropriate assumptions it is possible to find evidence that Raising needs to apply in NPs whose head nouns are not

nominalizations and hence cannot come from sentences. The syntax of *chance* and *opportunity* closely parallels that of *tendency*, as Postal notes (1974:354-5); Raising thus offers a way to transformationally relate such pairs of phrases as:

(1.1-23) (a) an opportunity for me to escape:
             my opportunity to escape

          (b) a chance for me to escape:
             my chance to escape

Further, *role* (as in *Murphy's role as police chief*) may be susceptible to an analysis similar to that proposed for *continuation*.

Second, and more interestingly, sentences such as the following present difficulties for the Raising analysis:[5]

(1.1-24) (a) John has a tendency to sleep late.

          (b) John is characterized by a tendency to sleep late.

          (c) A tendency to sleep late is John's most notable attribute.

          (d) Of all John's attributes, the most striking is an egregious tendency to sleep late.

Consider how we might account for the fact that *John* is semantically the subject of *sleep late* in all of these. Can we do this by generating *John* as deep structure subject of *sleep late*? If we adopt such an approach, then Raising will get us from *the tendency for John to sleep late* to *John's tendency to sleep late*; but what then? It might be possible to get (a) from *John has John's tendency to sleep late* by some sort of Equi, and such an analysis might even gain independent support from the fact that *have* and the possessive convey similar meanings; but (b), (c), and (d) would introduce great complications into such a treatment.

However, the problem can be attacked from a different angle. The noun *tendency* (like *beauty*, *sloth*, *chutzpah*, and others) signifies an attribute, and whenever such a noun is used in a sentence, some part of the grammar--presumably the semantic interpretation rules--has to be able to determine who or what the attribute is being ascribed to. Given this, all we need is to specify that whatever is interpreted as the subject (if we may call it that) of *tendency* is also to be interpreted as the subject of the VP that follows *tendency*. That is, it seems feasible to let semantic interpretation rules do the work of Raising, at least in this case, and perhaps in all cases (over verbs as well as nominalizations).

This argument depends, of course, on the plausibility of the claim that assigning attributes to referents is the job of semantic rather than syntactic rules--which may not be, in itself, a testable claim. The inquiry has reached the stage at which metatheoretical considerations and opinions about heuristics play at least as large a role as data--and this is not the place to carry on the generative semantics-interpretive semantics debate. Either way, however, Postal's arguments about Raising provide no real evidence against the Lexicalist Hypothesis, for even if a syntactic rule of Raising does apply to *tendency* and the like, there is no evidence that it precedes Nominalization.

1.1.2 NP structure

Chomsky's next set of arguments is directed at showing that gerundive nominals occur in syntactic structures similar to sentences, while derived nominals occur in phrases with the structure of non-sentential NPs. (Chomsky simultaneously urges the logically distinct claim that the deep structure of noun phrases is in various ways closer to surface structure than had previously been supposed; however, this second hypothesis does not interact with the Lexicalist Hypothesis per se, and I shall not review the arguments for it in detail.)

To begin with, gerundive nominals take aspect markers (*have...en*, and perhaps *be...ing*), manner adverbs, and negation; derived nominals do not.

(1.1-25)   Perfect aspect:

    (a)   John has proved the theorem.

    (b)   John's having proved the theorem

    (c)   *John's have proof of the theorem

(1.1-26)   Progressive aspect:

    (a)   John is criticizing the book.

    (b)   ?John's being criticizing the book

    (c)   *John's be criticism of the book

(1.1-27)   Manner adverbials:

    (a)   John proved the theorem ingeniously.

    (b)   John's proving the theorem ingeniously

    (c)   *John's proof of the theorem ingeniously

(1.1-28) Negation:

  (a) John did not prove the theorem.

  (b) John's not proving the theorem

  (c) *John's not proof of the theorem

A few notes on these examples are in order. The question mark on (1.1-26) seems to follow from a general constraint against two -*ing*'s in a row: ?*quitting smoking*, ?*keeping looking*. Note that *John's non-proof of the theorem* does not have the right meaning to fit into the (1.1-28c) slot; using *non-proof* entails that John produced something that was not a proof, while (1.1-28a, b) do not entail that John produced anything at all. Finally, the absence of manner adverbs in derived nominals could of course be explained by having Nominalization convert *John proved the theorem ingeniously* into *John's ingenious proof of the theorem*; we will consider this approach presently.

The behavior of gerundives here is precisely that predicted by the uncontroversial analysis that derives them from sentences in which TENSE + (MODAL) (the first part of the Aux) is replaced by a special gerundive marker. Tense and Modal are the only sentence elements that are missing from gerundives.

Derived nominals, on the other hand, have some noun-like properties that gerundives lack. They pluralize and quantify:[6]

(1.1-29) (a) some of John's proofs of the theorem

          (b) *some of John's proving(s) the theorem

And where gerundives have adverbs, derived nominals regularly have adjectives:

(1.1-30) (a) John's criticizing the book stupidly

          (b) John's stupid criticism of the book

(1.1-31) (a) his walking erratically through the park

          (b) his erratic walk through the park

Most strikingly, the 'object' of a derived nominal has to be preceded by a preposition (usually *of*), giving the phrase a structure like that of many NPs that are not nominalizations:

(1.1-32) (a) John's clever proof of the theorem

          (b) John's red book of matches

          (c) Asimov's latest anthology of science fiction

It is surely implausible to derive *book of matches* by nominalization from an abstract verb (meaning what?). In fact, no one doubts that Possessor - Adj - N - PP structures have a source other than nominal-ization, and as long as this is so, the lexicalist prediction that derived nominals occur in non-sentential NP structures holds up. The reason Chomsky devotes so much attention to these NPs in "Remarks on Nominalization" is that he is also arguing against various relative clause reduction analyses of them that have been proposed; this is, however, a logically independent issue.

One other point needs attention. At first sight, the pairings below seem to provide an argument for a nominalization transformation:

(1.1-33) (a) John refused suddenly.

(b) John's sudden refusal

(1.1-34) (a) Grendel arrived noisily.

(b) Grendel's noisy arrival

(1.1-35) (a) Holbein painted skillfully.

(b) Holbein's skillful painting

It looks as though the adjectives in the (b) examples are derived transformationally from the adverbs in the (a) examples, and thus that nominalization is a transformation. Chomsky (1970:195) presents his treatment of such evidence in the form of a dilemma. (His argument is brief, and I am expanding on it somewhat.)

Suppose the transformational analysis is accepted. There are nominalizations that contain adjectives that have no adverb counter-parts in the corresponding sentences, for example:

(1.1-36) (a) John's uncanny resemblance to Bill

*John resembled Bill uncannily.

(b) Lincoln's untimely death

*Lincoln died untimely.

(c) Admiral Byrd's polar explorations

*Admiral Byrd explored polarly.

Suppose we derive *uncanny*, *untimely*, and *polar* in the examples above from adverbs that are required to undergo the rule that converts them into adjectives in connection with nominalization, and cannot appear unconverted. Then, in addition to these specially marked adverbs, the lexicon will also have to contain entries for *uncanny*, *untimely*,

and *polar* as adjectives, to account for their occurrence in non-nominalized phrases like *uncanny phenomena, untimely misfortunes,* and *polar flora and fauna.* Having done this, we need a full-fledged lexicalist apparatus of redundancy rules and the like to relate the adverbs to the adjectives.

The alternative is to treat *uncanny, untimely, polar,* and the like purely as adjectives, and posit that they get into nominalizations the same way they get into ordinary NPs. But if we do this, then the adjective-adverb correspondences noted above provide no argument for a nominalization transformation.

Ross (1973) challenges Chomsky's generalization by pointing out that not only derived nominals, but also gerundives, can take *this, that,* and a restricted range of quantifiers:

(1.1-37) This looking all the addresses up has got to stop.

No getting arrested will be tolerated.

All that looking up addresses got to me.

He takes this as showing that gerundives are not as devoid of NP structure as Chomsky might claim, and he considers it evidence for his hypothesis that gerundives and derived nominals are intermediate points on a continuum from S to NP (with gerundives closer to S and derived nominals closer to NP). However, Ross's observations do not in fact substantially weaken Chomsky's argument. After all, according to Ross, derived nominals do pattern exactly like morphologically simple nouns (1973:170); and conceding determiners and quantifiers to gerundives does not amount to much of a concession. After all, on Chomsky's analysis a gerundive is an S embedded in an NP; the quantifier can perfectly well belong to the part of the NP that is outside the S.

Chomsky's concern in showing that nominalizations have the structure of NPs is to point out that this fact is consistent with the Lexicalist Hypothesis. However, it is also consistent with a suitably constrained transformational treatment of nominalization. In section 1.1.1 we saw that Nominalization (if it is a transformation) must precede, as far as we can tell, all other transformations. The facts just discussed show that it must also be constrained to produce NP-like structures as its output.

Such a constraint has a familiar ring to it; it is in fact a special case of Emonds' (1970) Structure Preserving Hypothesis, which constrains ALL transformations to produce base-generable output phrase structures, except for local transformations (those that operate on at most two adjacent elements at least one of which is not phrasal) and root transformations (those that move, copy, or insert nodes that end up immediately dominated by the highest S in the tree).

Nominalization is neither a root nor a local transformation; the constraint that it must produce NP structures as output thus follows from Emonds' more general hypothesis. The transformation that forms gerundive nominals from structures like (1.1-1) above :is plausibly treated as a local operation that changes *NP ing* to *NP's ing* (after which Affix Hopping attaches the *ing* to the verb); it is thus not required to be structure preserving.

So far, then, the evidence presented has been consistent both with the Lexicalist Hypothesis and with a nominalization transformation that is precyclic and obeys the Structure Preserving Hypothesis. What is needed is a way to discriminate between the two.

1.1.3 Derivational idiosyncrasy

Chomsky (1970:189) states the argument from derivational idiosyncrasy as follows:

> The idiosyncratic character of the relation between the derived nominal and the associated verb has been so often remarked that discussion is superfluous. Consider, for example, such nominals as *laughter, marriage, construction, actions, activities, revolution, belief, doubt, conversion,* ...and so on, with their individual ranges of meaning and varied semantic relations to the base forms. There are a few subregularities that have frequently been noted, but the range of variation and its rather accidental character are typical of lexical structure. To accomodate these facts within the transformational approach (assuming, as above, that it is the grammatical relations in the deep structure that determine meaning) it is necessary to resort to the artifice of assigning a range of meanings to the base form, stipulating that with certain semantic features the form must nominalize and with others it cannot. Furthermore, the appeal to this highly unsatisfactory device, which reduces the hypothesis that transformations do not have semantic content to near vacuity, would have to be quite extensive.

The point is that various properties of derived nominals, most strikingly their meanings, are not predictable by rule from the properties of the verbs from which they are derived. For instance, a book can go through two *editions*, one paperbound and one hard-cover, without properly being said to have been *edited* twice, and one can suffer an *attack* of asthma without being *attacked* by it.

As Chomsky points out, this kind of idiosyncrasy can be handled within the transformational approach by giving *edit* (for example) an additional sense that corresponds exactly to *edition* and is marked to

obligatorily nominalize. The need to subdivide lexical entries to handle separate but related senses has been evident at least since Katz and Fodor (1963), though there has long been disagreement as to exactly how to do it.

Adding such a provision for obligatorily nominalized senses of verbs does not, as Chomsky claims, "reduce the hypothesis that transformations do not have semantic content to near vacuity." Rather, it predicts that nominalizations and the verbs from which they are derived are either synonymous (modulo the portion of the meaning attributable to the environment of nominalization, such as the suffix), or else different IN THE WAY THAT SENSES OF LEXICAL ITEMS DIFFER FROM EACH OTHER, i.e., in more or less haphazard ways due to the accidents of etymology. There is still not supposed to be any regular difference of meaning attributable to the transformation. This constitutes an interesting, and testable, claim; in fact, it seems to be correct.[7]

But if we adopt this approach, we are left with no explanation of why there are no idiosyncratic gerundives. If both gerundive nominals and derived nominals are formed transformationally, and if extra senses can be introduced and marked to undergo transformations obligatorily, then gerundives ought to be able to behave in the same way as derived nominals. They don't. We have run into a qualitative difference between gerundives and derived nominals that cannot be accounted for by an appeal to a familiar device such as rule ordering and/or the Structure Preserving Hypothesis. In this regard the Lexicalist Hypothesis does appear preferable to a transformational account of derivation.

1.2 Post-syntactic incorporation in Greenlandic

Logically, if not chronologically, the next thing to be considered is Sadock's (1980) argument that certain incorporation phenomena in Greenlandic Eskimo, traditionally described as derivational, cannot be handled in the lexicon because they are ordered after syntactic transformations. Sadock describes several types of incorporation, but I shall focus on the one that provides the strongest argument.

This is what Sadock calls object incorporation. Actually, as he points out, INCORPORATION is somewhat of a misnomer, since what happens is that verbs are formed from nouns by adding verb-forming suffixes that never occur alone, not by compounding them with verbs proper. For instance:[8]

(1.2-1) qimmeq
         'dog'

         qimmeqarpoq
         'dog-have-INDIC.-3SG.'
         'He has a dog.'

(1.2-2) sapangaq
         'bead'

         sapangarsivoq
         'bead-buy-INDIC.-3SG.'
         'He bought beads.'

It is, however, possible to attach the verb-forming suffixes to an
"empty stem" *pi*, glossed as 'thing':

(1.2-3) peqarpoq
         'thing-have-INDIC.-3SG.'
         'He has something.'

(1.2-4) pisivoq
         'thing-buy-INDIC.-3SG.'
         'He bought something.'

When this is done, it is possible to express an overt object:

(1.2-5) qimmimik peqarpoq
         'dog-INST. thing-have-INDIC.-3SG.'
         'He has a dog.'

(1.2-6) sapanngamik pisivoq
         'bead-INST. thing-buy-INDIC.-3SG.'
         'He bought a bead.'

These are instances of the Greenlandic antipassive construction, in
which the verb is morphologically intransitive but semantically tran-
sitive, and the notional object is in the instrumental case. "Incor-
porated" verbs like these are obligatorily antipassive.

The next portion of Sadock's argumentation is worth quoting in
extenso (with his example numbers):

> A striking fact about incorporated objects, and one that argues
> powerfully for deriving them syntactically rather than morpho-
> logically [i.e., lexically], is that they may be modified. The
> modifier appears as a separate word in the instrumental case--
> the same case that it would have if it were the modifier of a
> non-incorporated object of a free-standing, formally intransitive
> verb. Compare 27, in which the object is incorporated, with the
> synonymous 26, in which it is not:
>
>   (26) Sapanngamik kusanartumik   pisivoq.
>   'bead-INST. beautiful-NOM.-INST. thing-get-INDIC.-3SG.'
>   'He bought a beautiful bead.'

> (27) Kusanartumik   sapangarsivoq.
> 'beautiful-NOM.-INST. bead-get-INDIC.-3SG.'

If 27 were derived either from a structure very much like 26,
or from a somewhat more abstract structure into which the
empty stem had not yet been inserted, then the case of the
modifier would automatically be assigned by independently-
needed rules.

However, the generalization that is so obvious between
26 and 27 would be obscured if object-incorporating verbs had
to appear fully formed in deep structure. I assume that
direct objects of formally intransitive, notionally tran-
sitive verbs have instrumental case assigned to them by syn-
tactic rule. I further assume that a syntactic rule copies
the case of a head noun onto all its non-possessive modifiers.
Now suppose that verbs like *sapangarsivoq* in 27 are lexically
rather than syntactically derived. Then, if modifiers of
incorporated objects were treated as syntactically distinct
from direct objects, an ad-hoc syntactic rule would be needed
to assign instrumental case to them. And if such modifiers
were treated as syntactically identical to direct objects,
an ad-hoc semantic rule would be needed to interpret them as
modifiers. Either way, a generalization would be lost. (Sadock
1980:307-8)

As if this were not enough, note that in examples like (1.2-1), the
incorporated noun stem is unspecified with regard to plurality;
qimmeqarpoq could equally well mean 'He has dogs.' But when a modi-
fier is present, the number of the modifier determines the understood
number of the noun. Contrast Sadock's (27), cited above with

> (1.2-7) kusanartunik   sapangarsivoq.
> 'beautiful-NOM.-PL.-INST. bead-get-INDIC.-3SG.'
> 'He bought beautiful beads.'

The former is unambiguously singular, while the latter is unambiguously
plural. If the incorporation is performed by a process that follows
case and number agreement and that removes the number marking from the
noun stem, this is no problem. If, on the other hand, incorporation
is done in the lexicon, the association of the number of the modifier
with the noun stem will have to be done by a semantic interpretation
rule, otherwise unmotivated. And this semantic rule will have to deal
with the fact that wherever grammatical and semantic number are at odds
with each other (a common situation in Greenlandic), the modifier agrees
with the grammatical rather than the semantic number of the noun stem.
For example, the word for 'sled' is grammatically plural and, even when
incorporated, requires plural agreement:

(1.2-8)   angisuunik   qamuteqarpoq
          'big-NOM.-PL.-INST. sled-have-INDIC.-3SG.'
          'He has a big sled.'

(1.2-9)   *angisuumik   qamuteqarpoq
          'big-NOM.-INST.   sled-have-INDIC.-3SG.'

It is even possible to say 'He has one sled' in such a construction
by adding the morphologically plural form of the word for 'one'.

Obviously, a lexical-interpretive treatment of these data misses
a number of generalizations; the case for a syntactic treatment is
strong. But there is a difficulty.   Implicitly, Chomsky (1970) wants
to use the lexicalist treatment not just for derived nominals, but
for the whole of derivational morphology; this approach is made ex-
plicit by Jackendoff (1972). Traditional criteria for distinguishing
inflection from derivation include the following:

(a) Derivation produces forms that undergo inflection, but not
     (usually) the other way around.

(b) Derivation, but not inflection, can derive one part of
     speech from another. (Participles, which are traditionally
     inflectional, present a problem with this criterion.)

(c) The existence of any member of an inflectional paradigm
     generally implies the existence of the complete paradigm;
     derivation, on the other hand, is commonly subject to
     accidental gaps and incomplete productivity.

(d) Derivation, but not inflection, can produce unpredictable
     changes of meaning (cf.*edit*:*edition* above).

Noun incorporation fulfills the first two criteria: it forms
verbs from noun stems, and these then undergo normal verb inflection.
So it is derivational, and (according to the Lexicalist Hypothesis)
has to be handled in the lexicon. But it follows various syntactic
agreement transformations, which in turn follow transformations like
Passive and Antipassive, and therefore cannot be handled in the lexicon.
Is there a way out of this paradox?

For an answer, remember that the strongest argument for lexicalism
was based on derivational idiosyncrasy, corresponding to criteria (c)
and (d) above. But Greenlandic noun incorporation is, as far as I can
determine, not idiosyncratic at all (Sadock 1980; Rischel 1971). If
this is indeed the case, then the Lexicalist Hypothesis can be saved
by redefining its scope; LEXICAL TREATMENTS ARE REQUIRED, NOT OF ALL
TRADITIONALLY DERIVATIONAL PROCESSES, BUT ONLY OF THOSE THAT SHOW
IDIOSYNCRASY.

Though there are serious infelicities in his statement of the hypothesis, this seems to be essentially what Sadock proposes. He concludes that universal grammar must allow "pre-affixal syntax", which he defines as follows (1980:301-2):

> First, pre-inflectional syntax is allowed... In addition, meaning-bearing morphemes may be rearranged by syntactic rules into word-sized constituents. To put teeth into this theory, suppletion must be limited to inflection; thus suppletion of both roots and derivational affixes would be forbidden. Under such a proposal, *teacher* could be derived from an abstract source in which the verb *teach* and the content of the agentive suffix were not continuous; but *redden* and *Americanize* would have to be lexically derived, because of the suppletive relation between the causative morphemes *-en* and *-ize*.

Sadock seems to equate idiosyncrasy with suppletion (the impossibility of predicting morphology from syntax and semantics) and ignores the impossibility of predicting derived semantics from root semantics in cases like *edit:edition*. However, his proposal seems to be on basically the right track.

Note that, since Greenlandic noun incorporation forms verbs that begin with noun stems, it is also necessary either to modify Jackendoff's (1972:13) claim that transformations cannot change category labels, or to analyze noun incorporation in such a way that a transformation merges a noun with a verb, rather than (strictly speaking) turning a noun into a verb. The latter is probably preferable (see Rischel 1971).

1.3 The passive paradox

Wasow (1977) shows that, within a lexicalist framework, there are good arguments that Passive must be handled lexically, and also good arguments that it must be handled syntactically, in English.

The arguments that Passive is lexical fall into two sets. First, many passive verb forms show adjectival syntactic properties; and if, as Jackendoff (1972) and others have hypothesized, transformations cannot change verbs into adjectives, then Passive has to be handled lexically. Passive verb forms behave like adjectives in that they appear in prenominal modifying position, appear as complements to verbs like *seem* and *act*, and take as an intensifier *very* (as opposed to *very much*):

(1.3-1) a $\left\{\begin{array}{l}\text{green}\\\text{new}\\\text{discarded}\\\text{thrown-away}\\\text{crushed}\end{array}\right\}$ cardboard box

(1.3-2) John $\begin{Bmatrix} \text{seemed} \\ \text{acted} \end{Bmatrix}$ $\begin{Bmatrix} \text{happy.} \\ \text{surprised by the news.} \\ \text{annoyed at us.} \end{Bmatrix}$

(1.3-3) John is very $\begin{Bmatrix} \text{new.} \\ \text{respected.} \\ {}^{?}\text{looked up to.} \end{Bmatrix}$

Second, many passives can take the prefix *un-* that attaches to adjectives to signify 'not...': *unerased, unmarked, unseen, untouched.* (This is distinct from the *un-* that attaches to change-of-state verbs in all tenses and voices to signify reversal, as in *untie, unlock,* and the like.)[9] Since the prefixation of *un-* is derivational and therefore lexical, and since Passive precedes it, Passive has to be lexical.

But there is a caveat. In view of the Greenlandic evidence above, the fact that *un-* is derivational does not show that it is lexical. The crucial criterion is idiosyncrasy, and idiosyncrasies involving *un-* are hard to find; the sense of *uncounted* 'uncountable' in *uncounted hosts* may be an example. But if *un-* is not lexical, there is no reason to suppose that passivization preceding *un-* is lexical either. This, then, is a potential weak spot in the argument, but since Wasow did not know about the Greenlandic evidence and did not make the modification in lexicalist theory that I have suggested on the base of it, I shall not pursue this objection here.

On the other hand, Passive crucially follows Dative Movement, which is generally thought to be syntactic, in sentences like *The turtle was given an ear of corn,* and no lexical rule can follow a syntactic transformation. (This fact was noted by Freidin 1975 as an objection to his hypothesis that Passive is lexical.) Further, if passive participles are adjectives, they are the only adjectives that can take noun complements:

(1.3-4) John was $\begin{Bmatrix} \text{elected} \\ {}^{*}\text{happy} \end{Bmatrix}$ President.

John is $\begin{Bmatrix} \text{considered} \\ {}^{*}\text{obvious} \end{Bmatrix}$ a fool.

There are basically three ways out of the resulting paradox:

(a) Split Passive into two rules, one lexical and the other syntactic, and argue that they function differently. This is Wasow's approach; I shall discuss it below.

(b) Treat Passive as solely syntactic. This entails repudi-
     ating the arguments that Passive is lexical; to be specific,
     verbs have to be turned into adjectives by a syntactic
     transformation, and post-passive *un-* has to be attached
     transformationally. This is tantamount to rejecting the
     Lexicalist Hypothesis.

(c) Treat not only Passive, but also Dative Movement and
     whatever else it crucially follows, as lexical rules.
     This approach results in a theory similar to that of
     Bresnan (1978), in which all relation-changing rules
     are lexical. (Bresnan in fact has two Passive rules,
     corresponding to Wasow's two rules except that both
     are lexical.) Note that many of Chomsky's original
     arguments for lexicalism are inapplicable to such a
     theory, since they depend on showing a contrast between
     derivation and other syntactic processes, and a theory
     that puts all relation-changing rules in the lexicon
     denies such a contrast.

Wasow argues that splitting Passive into two rules is the correct
move because passives that have to be syntactically derived (because
they follow Dative Movement or other transformations) do not show the
Lexical Passive properties of being adjectives and taking *un-*; while
the passives that have to be lexical do not correspond precisely to
the output of the syntactic Passive rule.

That passives that follow Dative Movement do not have adjectival
syntax is shown by the following:

(1.3-5) (a) The man was ⎧spattered with paint.⎫
                           ⎨told the story.   ⎬
                           ⎩given five dollars.   ⎭

         (b) The man ⎧looks⎫   ⎧?spattered with paint.⎫
                       ⎩seems⎭   ⎨*told the story.   ⎬
                                 ⎩*given five dollars.   ⎭

         (c) the ⎧spattered⎫ man
                   ⎨*told   ⎬
                   ⎩*given   ⎭

There is a clear contrast in (b), even if none of the sentences is
perfectly acceptable.

Passives that follow Dative Movement do not take *un-*: *Bill was
untold the story*. Neither do passives that follow Raising: *John is
known/*unknown to be a Communist*.

On the other hand, the properties of passives for which there is a good case for a lexical treatment--i.e., those that have adjectival syntax and take *un-* --do not correspond precisely with the properties predicted by the syntactic Passive rule. Wasow notes that adjectival passives can require adverbs that are not required by the verbs from which they are derived:

(1.3-6) (a) *two noticed errors
              two widely noticed errors

         (b) *this read book
              this rarely read book

         (c) *This chicken smells killed.
              This chicken smells freshly killed.

A stronger argument comes from Bresnan (1978:8), who points out that there are various lexical passives of intransitive verbs. An *untraveled* person, for instance, is a person who has not *traveled*, not someone who has not *been traveled*. The situation is similar with *unhurried*, *unpracticed*, and perhaps other adjectives. Apparently, the lexical passive rule, but not the syntactic rule, can apply to intransitive verbs.[10]

To summarize, Wasow argues convincingly that the superficially implausible hypothesis that English has two separate passive rules (the only analysis consistent with both the Lexicalist Hypothesis and the data) is in fact well motivated on other grounds.

Nanni (1980) gives a parallel argument that adjective phrases like *easy to please* must have both transformational and lexical sources. In addition to the standard derivation of, for example, *John is easy to please* from either [*To please John*] *is easy* (by *Tough*-Movement) or *John is easy* [*to please John*] (by Equi-object deletion), she proposes a lexical rule that creates [Adj *to* V]$_{Adj}$ structures. In the interest of brevity I shall not review her argument in detail; the gist of it is as follows.

In general, adjectives that are followed by any kind of complement cannot appear in prenominal position:

(1.3-7) a difficult book
         *a difficult for us to read book
         *a difficult to read aloud book
         *a difficult to try to get Bill to read book
         *a more difficult than ever book
         *a difficult for young readers book

However, in this regard adjectives followed solely by *to V* act like the complement-less adjectives:

(1.3-8)   a difficult-to-read book
          a hard-to-find manuscript
          the hard-to-understand part of the theorem

This suggests that *Adj to V* constructions are structurally different from the other instances of adjectives with complements, many of which (such as the first three starred phrases of (1.3-7) above) are transformationally derived by *Tough*-Movement or its equivalent. In fact, if one accepts the argument of Chomsky (1970) and others that pre-nominal adjectives are base-generated in place, then the data require either an otherwise unmotivated adjective-preposing rule or a lexical source of $[\text{Adj } to \text{ V}]_{\text{Adj}}$ structures.

As further evidence that the required new rule is lexical, Nanni points out data like the following:[11]

(1.3-9)   an easy-looking pattern
          an easy-to-make-looking pattern
          *an easy-to-make-quickly-looking pattern
          *an easy-for-us-to-make-looking pattern (etc.)

It looks as though the infinitive gets attached to the adjective before *looking* does; and if the rule that does the latter is lexical, then so is the former, which precedes it.

If Nanni's argumentation is valid, then a pattern is beginning to emerge: at least two English transformational rules have similar but not identical counterparts among the lexical rules. This suggests two interesting conjectures:

(1) Perhaps there is something fundamentally wrong with the framework and method used in constructing these arguments; perhaps all grammatical processes in a certain class (for instance, relation-changing transformations) will seem double, as an artifact of the method of analysis. This question urgently needs investigating.

(2) Alternatively, perhaps the duplication of function between these transformational rules and lexical rules is psychologically real and is a result of diachronic reanalyses that have not completely gone through. The syntactic rules of a language change when reanalysis occurs, i.e., when children learning the language internalize different rules for handling particular constructions than the previous generation did, and are able to get away with doing so. Perhaps the lexical (or transformational) Passive rule, for example, represents a reanalysis that speakers have made at an early stage of language acquisition, and

that has proved inadequate for dealing with sentences subsequently encountered, making it necessary to add the other, transformational (or lexical) rule alongside the first. If this is so, then the kind of evidence that Wasow and Nanni cite is available only because vast numbers of speakers have carried out the reanalysis in parallel fashion.

This failed-reanalysis conjecture of course suggests a wide range of empirical tests. It would be extremely suggestive if the speech of children in a particular age range provided motivation for only one of the two rules in each lexical-transformational pair.

2. The operation of lexical rules

So far, we have been considering the Lexicalist Hypothesis in its minimal form, i.e., the claim that derivational relationships (or a subclass of derivational relationships not including Greenlandic noun incorporation) are expressed in the lexicon, not in the trans- formational component. Very little has been said about lexical rules except that they are ordered before all transformations and differ from them in some unspecified way. So the search for evidence for lexicalism has amounted to a search for pre-transformational non- transformations. Considered as a theory, this minimal form of lexi- calism has relatively little to say about language; considered as a framework, however, it opens up a new range of possible theoretical claims. The remainder of this paper is concerned with arguments for lexicalism that rest on new theoretical claims that can only (or only easily) be stated in a lexicalist framework.

2.1 Two conceptions of lexical rules

The function of lexical redundancy rules (henceforth LRRs) is to express regularities that hold across distinct lexical entries; to show, for instance, that the relationship between *decide* and *decision* is roughly the same as that between *convert* and *conversion*. This is true whether the lexical entries of *decide* and *decision* form parts of a single branching structure (as Chomsky originally envisioned) or are completely separate. Chomsky (1970) elaborates neither the branching structures nor the lexical redundancy rules; Jackendoff (1975) elaborates the LRRs, with the implicit result that the branching struc- tures are not needed: if *decide* and *decision* are related by rule, they do not need to be related in any other way.

Perhaps the most obvious theory of LRRs is what Jackendoff calls the IMPOVERISHED-ENTRY THEORY. (In what follows, I follow Jackendoff (1975); however, for convenience and since I have nothing to say about phonology, I shall use ordinary spellings in place of phonological representations.) In the impoverished-entry theory (IET), lexical entries fit into the following schema:

(2.1-1)

$$\begin{bmatrix} \text{entry number} \\ \text{/phonological representation/} \\ \text{syntactic features} \\ \text{SEMANTIC REPRESENTATION} \end{bmatrix}$$

The entry number is merely an index that allows a lexical entry to be referred to without quoting part of all of its contents; the entry numbers in the following examples will be purely arbitrary. For instance, the lexical entry for *decide* looks something like this:

(2.1-2)

$$\begin{bmatrix} 784 \\ \textit{decide} \\ +V \\ +[NP_1 \ \_ \ \text{on} \ NP_2] \\ NP_1 \ \text{DECIDE ON} \ NP_2 \end{bmatrix}$$

The semantic representation here is of course a fudge, adequate for the present purpose but not by any means adequate as a description of the semantics of *decide*.

The LRR that relates *-ion* derivatives to their roots looks like this:

(2.1-3)

$$\begin{bmatrix} X \\ Y + \textit{-ion} \\ +N \\ +[NP_1\text{'s} \ \_ \ (P) \ NP_2] \\ \text{ABSTRACT RESULT OF ACT} \\ \quad \text{OF } NP_1\text{'S } z\text{-ING } NP_2 \end{bmatrix} \longleftrightarrow \begin{bmatrix} W \\ Y \\ +V \\ +[NP_1 \ \_ \ (P) \ NP_2] \\ NP_1 \ z \ NP_2 \end{bmatrix}$$

Here W, X, and Z are variables that stand for (respectively) the two entry numbers, the phonological shape of the root, and the semantic representation of the root. If there is no preposition

in the subcategorization of the root verb (as in *convert*, which
does not take a preposition before its notional object), then
there is no preposition in the subcategorization of the *-ion*
derivative either, and a syntactic transformation inserts *of*.
Phonological rules change *decide+ion* into *decision*, and *convert+*
ion into *conversion*. (Hust (1977.85) has a number of criticisms
of Jackendoff's formulation of this rule and of various lexical
entries; however, they do not directly affect the theoretical
claim Jackendoff is making.)

The lexical entry for *decision* under the IET then looks
something like this:

$$(2.1-4) \quad \begin{bmatrix} 893 \\ \\ \text{derived from 784} \\ \quad \text{by rule } (2.1-3) \end{bmatrix}$$

That is, it is only necessary to identify the root and the rule by
which the derivation is performed, if the derivative has no idio-
syncrasies. If *decision* had properties that are not predictable
by rule, then they would be listed in lexical entry 893 also.

But another conception of the form and function of lexical
redundancy rules is possible. Consider what it means to say that
two lexical entries are related. A lexical entry is related to
another if some of the information in one can be predicted by rule
from the other. (Here *rule* means 'regularity'; we do not formulate
rules that derive exactly one form and capture no generalization.)
For instance, *decision* is related to *decide* but not to *jelly*. The
concept of relatedness leads to a plausible psycholinguistic claim:
a new lexical item is easier to learn if it can be related by rule
to an existing lexical entry than if it cannot. For example,
imagine a speaker of English whose lexicon includes, among other
things, the *-ion* LRR above and the word *elude*. The claim is that
such a speaker will have an easier time learning a new word meaning
'abstract result of the act of eluding' if it is pronounced *elusion*
than if it is pronounced *glurp*.[12]

Lexical redundancy rules, then, serve to state that some of the
information in one lexical entry is predictable from another. In
the IET, the rules do this by deriving lexical entries from other
lexical entries. The lexicon (entries plus rules) consists solely
of unpredictable information. Thus, in the IET, as in most of gener-
ative grammar, the criterion for deciding what constitutes a signifi-
cant generalization is explicitly or implicitly hidden in the notation;
it is assumed that whatever gives a shorter description (expressed
in the language of brackets and parentheses) is a better, more

generalization-capturing grammar. This is, however, a rather arbitrary criterion. An alternate approach is possible: give *decide* and *decision* (for instance) complete, fully specified lexical entries, and use the rule TO STATE WHAT PARTS OF THEIR CONTENTS ARE PREDICTABLE. This is the full-entry theory (FET), and the burden of Jackendoff (1975) is to show that it is preferable to IET. (Of course, it is easy to construct full-entry and impoverished-entry theories that are notational variants of each other, so Jackendoff's arguments amount to arguments for deviating from this ambivalent theory in directions that are only possible in FET.)

In the full-entry theory, *decision* has the following lexical entry:

$$(2.1-5) \quad \begin{bmatrix} decide\text{+}ion \\ \\ \text{+N} \\ \\ \text{+[NP}_1\text{'s \_\_ on NP}_2] \\ \\ \text{ABSTRACT RESULT OF ACT OF} \\ \quad \text{NP}_1\text{'S DECIDING ON NP}_2 \end{bmatrix}$$

(Entry numbers are no longer needed.) Rule (2.1-3) is not an entry-building process, but a schema into which the lexical entries of *decide* and *decision* are fitted. Its effect is to state, "Whenever there is a pair of lexical entries such that one fits into each side of this rule, all the information specified in one side of the schema can be predicted from the other." Idiosyncrasies on either side that do not fit into the schema are not carried over to the other side. (Some additional working out of how the theory is to deal with less-than-perfect fits is obviously needed.)

The rule, then, is used in computing the INDEPENDENT INFORMATION CONTENT of the lexicon which is less than the total amount of information in all the lexical entries. (This contrasts with IET, in which the length of the description, including all lexical entries, is supposed to be its own information measure.) Suppose for example that, in the pair *decide-decision*, *decide* is taken as basic. Then the independent information content of the pair consists of:

(a) all the information in the lexical entry for *decide*;

(b) the information that *decision* exists;

(c) the information that (2.1-3) is the applicable rule; and

(d) any information in the entry for *decision* that is not predictable by rule.

The sum of (a), (b), (c), and (d) is presumably less than the amount of information in the two entries per se. A higher sum obtains if *decision* is taken as basic and *decide* is related to it by rule, since the entry for *decision* now has to be counted entire, and it contains more information than *decide*. This corresponds to the theorist's intuition that *decide* is more basic than *decision*; it is a convention of the FET that the sequence of rule applications that produces the lowest independent information content is to be used. (This may amount to a convention that, of the two lexical entries that fit into an LRR, the one containing the less information is taken as basic; that is, LRRs normally operate from the lower-information side to the higher.)

2.2 Arguments for the full-entry theory

2.2.1 Rootless derivatives

Nouns like *perdition*, *illusion*, and *conflagration* present a problem to any theory of derivation: they have the morphological, syntactic, and semantic properties of *-ion* derivatives, but the verbs from which they are putatively derived (**perdite, *illude, *conflagrate*) do not exist. Three ways of dealing with them suggest themselves.

The first is to simply list them in the lexicon as unrelated to anything else. But this ignores the fact that they have much in common with *-ion* derivatives of the more normal sort: they have the suffix *-ion*, are nouns, and in general refer to an abstract process or result.
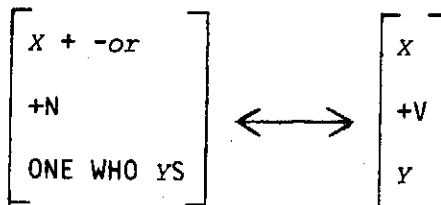
The second approach is to derive them in some way from hypothetical verbs **perdite*, **illude*, and **conflagrate*. In taking this approach the IET and the transformational account of nominalization are quite similar. In the IET, **perdite* is listed in the lexicon with a feature such as [-Lexical Insertion]; it can serve as the input to LRRs, but cannot itself be inserted into a base structure. In a grammar that handles *-ion* nominalization transformationally, **perdite* is marked as a positive absolute exception to the appropriate transformation, so that it has to be nominalized. These analyses make the claim that English would be simpler if *perdite* existed as a verb; the lexicon would be as it is now except that the [-Lexical Insertion] or exception feature on *perdite* would be deleted. In combintaion with the plausible assumption that there is functional pressure to reduce the information content of the lexicon, this analysis predicts that back formation should run rampant through the language, for back formation is nothing more than the loss of exception features, and the more back formation takes place, the simpler the language gets. But **perdite*, **illude* and **conflagrate* have not come into use, even after many years; this throws doubt on the analysis.

The third approach is to list, for example, *conflagration* (but not **conflagrate*) in the lexicon with a complete lexical entry, and specify in some way which parts of the entry are predictable (noun, abstract process or result, etc.) and which are not ('go up in flames'). Jackendoff incorporates this approach into the FET by allowing rules to apply rootlessly. For instance, *conflagration* fits into the left-hand (derivative) side of the schema,[13] but there is nothing to fit into the right-hand (root) side. This has the consequence that the information that is supposed to come from the root side of the schema cannot be counted as predictable, but the information that comes from the rule itself can. This seems to be the correct result. On this analysis, the independent information content of *conflagration* is less than it would be if the *-ion* rule were not applicable to it, but a back formation (*conflagrate*) would constitute an increase in the total information content of the lexicon, since a new entry would have to be established. (The increase would not be as great as that involved in adding a lexical entry not relatable to another by rule.) This analysis seems more consistent with actually observed back formation, which resembles a steady trickle rather than a flood.
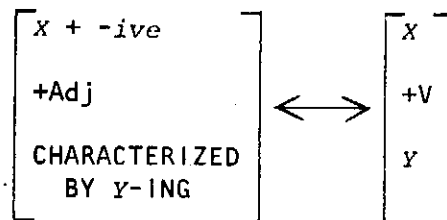
Since the third analysis can be set up only in the FET framework, the arguments for it constitute the arguments for FET.

Jackendoff constructs another argument, which strikes me as considerably weaker, based on sets of rootless derivatives, such as *aggressor, aggressive, aggression* (but not, at least in widespread use, **aggress*). The rules for the affixes (including the *-ion* rule, repeated here with some improvements) are, with various fudges, the following:

(2.2-1)   *-or* rule:

$$
\begin{bmatrix} X + -or \\ +N \\ \text{ONE WHO } YS \end{bmatrix} \longleftrightarrow \begin{bmatrix} X \\ +V \\ Y \end{bmatrix}
$$

(2.2-2)   *-ive* rule:

$$
\begin{bmatrix} X + -ive \\ +Adj \\ \text{CHARACTERIZED} \\ \text{BY } Y\text{-ING} \end{bmatrix} \longleftrightarrow \begin{bmatrix} X \\ +V \\ Y \end{bmatrix}
$$

(2.2-3) *-ion* rule:

$$\begin{bmatrix} X + \text{-}ion \\ +N \\ +[(NP_1 \text{ 's)} \_ ((P) \text{ } NP_2)] \\ \text{ABSTRACT RESULT OR} \\ \quad \text{PROCESS OF } NP_1\text{'S} \\ \quad Y\text{-ING } NP_2 \end{bmatrix} \longleftrightarrow \begin{bmatrix} X \\ +V \\ +[NP_1 \_ ((P) \text{ } NP_2)] \\ NP_1 \text{ } Y \text{ } NP_2 \end{bmatrix}$$
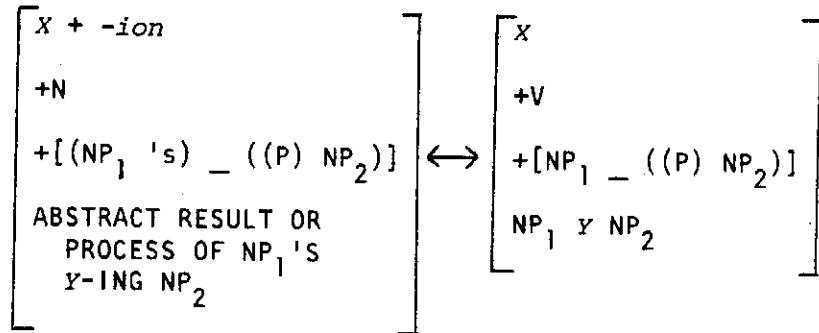
Consider first the relation between *aggressive* and *aggressor*, ignoring *aggression*. An argument has already been made for giving rootless derivatives lexical entries of their own and for not having a lexical entry for the nonexistent root. Thus, *aggressive* and *aggressor* are in the lexicon and have to be related in some way. Jackendoff proposes to relate them by stringing rules together. *Aggressive* fits the left side of the *-ive* rule; the right side of the *-ive* rule fits the right side of the *-or* rule; and by then applying the *-or* rule right to left we get *aggressor*. Rule-stringing thus makes it possible to relate forms that are derived from the same nonexistent root, without giving the root a lexical entry. And it doesn't matter whether *aggressive* or *aggressor* is taken as prior.

With the full triplet *aggressive-aggressor-aggression*, the possible orders of rule application are all the paths through the diagram below that connect all three items:

(2.2-4)

aggression

-ion rule

-or rule

-ive rule

aggressor
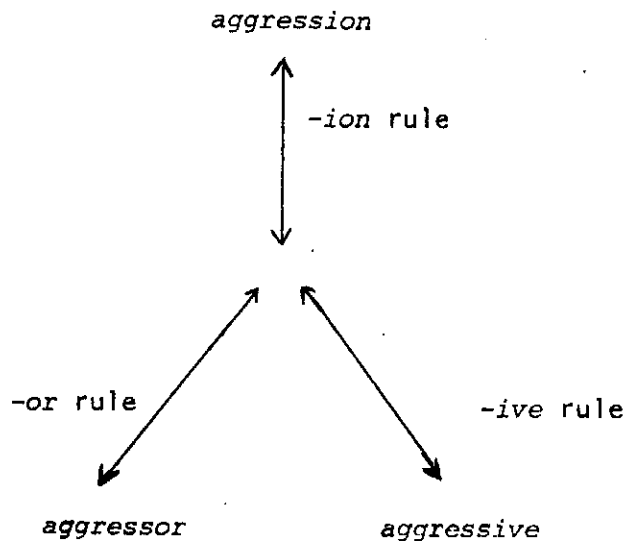
aggressive

That is, one can start at *aggression*, apply the *-ion* rule backward and the *-or* rule forward to get to *aggressor*, and then apply the *-or* rule backward and the *-ive* rule forward to get to *aggressive*; or one can go from *aggressor* to *aggression* and then to *aggressive*; and so forth. (*Aggress* has only a phantom-like existence as the junction point in the middle.)

There is a total of six possible paths, and the amount of information counted as predictable comes out about the same each way; this is supposed to reflect the fact that the language provides no motivation for choosing *aggression*, *aggressor*, or *aggressive* as prior to the other two members of the triple. But this is clearly an advantage only if one is eschewing psychological reality in order to describe language as a social phenomenon. The reason is that although it is clear that no member of the triple is prior in English AS A WHOLE, it is quite plausible that one of them is prior IN EACH SPEAKER'S HEAD--the one that that particular speaker learned first. (If the three words are more or less equally common, then so are the six orders of acquisition, which is why none of them is to be preferred to the others in describing the speech community as a whole.) This conjecture is probably open to experimental testing, and since Jackendoff elsewhere claims (or at least conjectures) a high degree of psychological reality for his theory, I presume he would consider this a potential objection to his argument.

2.2.2 Derivational idiosyncrasy

In the IET, the lexical entry for a derivative consists of a cross-reference to the root and the rule used to perform the derivation, plus any semantic features that the rule does not supply. For instance, the entry for *transformer* is something like the following (doing without entry numbers and rule numbers):
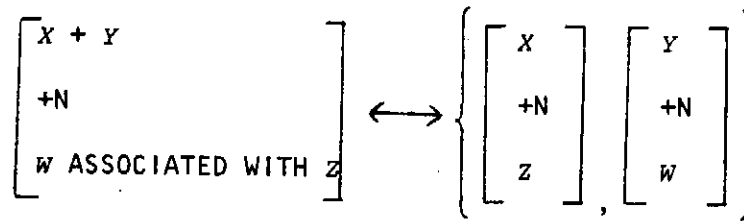
(2.2-5)
$$
\begin{bmatrix}
\text{derived from } transform \\
\quad \text{by } \text{-}er \text{ rule} \\
\text{...VOLTAGE OF ALTERNATING CURRENT...}
\end{bmatrix}
$$

The semantic markers for which VOLTAGE OF ALTERNATING CURRENT is a fudge have to be inserted in the semantic structure in the right place by some still un-worked-out mechanism.

That is how semantic features are added; but what if they have to be subtracted? That is, what if the derivative lacks semantic material present in the root and carried over by the rule? Jackendoff presents as particularly striking instances of this the nouns *strawberry* and *gooseberry*, which are plausibly analyzed as noun compounds but whose referents have nothing to do with straw or geese. The appropriate noun compound rule is something like the following (where ASSOCIATED WITH is a fudge for a variety of semantic relationships):

(2.2-6) Noun compound rule:

$$
\begin{bmatrix} X + Y \\ +N \\ W \text{ ASSOCIATED WITH } Z \end{bmatrix} \longleftrightarrow \left\{ \begin{bmatrix} X \\ +N \\ Z \end{bmatrix}, \begin{bmatrix} Y \\ +N \\ W \end{bmatrix} \right\}
$$

(Here the curly brackets stand for set membership, not disjunction.) This is the rule that accounts for such lexical items as *chicken coop*, *snowman*, and *sheepdog*.

If applied in the IET manner, this rule will build for *strawberry* a lexical entry showing the incorrect meaning BERRY ASSOCIATED WITH STRAW. The only way to prevent this from happening is to resort to some quite ad hoc mechanism, such as equipping the entry for *strawberry* with a set of minus features to cancel out all the plus features of the semantics of *straw* that will be brought in by the rule (assuming semantic features are binary, which is not necessarily a well-motivated assumption). This is clearly unsatisfactory.

The FET can, however, handle *strawberry* by refining the notion of a less-than-perfect fit of a lexical entry into a rule. Suppose that a morphological and syntactic fit is sufficient to allow the rule to be applied, even if the semantic material fits into the schema only partly or not at all. In this case, the noun compound rule applies, but is unable to use the semantics of *straw* to predict anything in the lexical entry of *strawberry*.

Some diachronic evidence that this is the correct analysis comes from the fact that *groseberry* has been reanalyzed as *gooseberry* (Jackendoff 1975:657), despite lack of semantic connection. If morphosyntactic similarity is indeed sufficient to invoke a lexical redundancy rule, then this reanalysis has effected a reduction in the independent information content of the lexicon, making the language slightly easier to learn.[14] Other examples are *bryd-guma* > *bridegroom* (unconnected to any earlier sense of *groom*), Ojibwa *otchek* > *woodchuck*, and dialectal reanalyses such as *cucumber* > *cow cumber* and *asparagus* > *sparrow grass* (Anttila 1972:92-3).
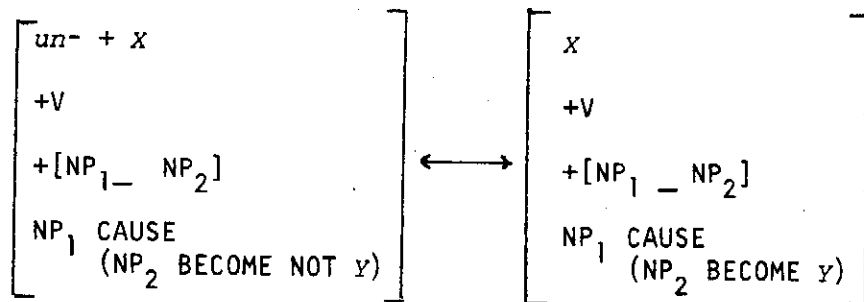
2.2.3 Redundant derivation

The arguments for FET reviewed so far have been Jackendoff's. However, there is another argument for full-entry lexicalism, based on the phenomenon of redundant derivation, in which an affix that normally serves to add a particular unit of meaning gets attached to a root whose meaning already includes that unit (Covington 1980).

English has a prefix *un-* that attaches to change-of-state verbs to form verbs that signify a reversal of this change of state, as in *lock:unlock, tie:untie, wind:unwind*. (This reversative *un-* is of course distinct from the *un-* that forms antonyms of adjectives, as already noted.)

The root verb to which *un-* attaches normally signifies putting something into a more marked or specialized state, and the derived *un-* verb signifies returning it to normal. Thus *unfreeze* is much less awkward than ?*unmelt*, since *freeze* is said of things whose normal state is not frozen (food, topsoil, and the like), while *melt* is said of things that may or may not be molten at normal temperatures (snow, ice, wax, metal). This fact can be expressed in fudged markerese by saying that the root verbs have meanings of the form CAUSE TO BECOME X and the derived *un-* verbs have meanings of the form CAUSE TO BECOME NOT X. The LRR for *un-* is then something like the following:

(2.2-7)   *un-* rule:

$$
\begin{bmatrix}
un\text{-} + X \\
+V \\
+[NP_{1\_} \quad NP_2] \\
NP_1 \ \text{CAUSE} \\
\quad (NP_2 \ \text{BECOME NOT } y)
\end{bmatrix}
\longleftrightarrow
\begin{bmatrix}
X \\
+V \\
+[NP_1 \_ NP_2] \\
NP_1 \ \text{CAUSE} \\
\quad (NP_2 \ \text{BECOME } y)
\end{bmatrix}
$$

That covers the ordinary, non-redundant use of *un-*. But there are also lexical items like the following (all from the *Oxford English Dictionary*):

unpick   'pick a lock'; later 'unstitch'

unrip   'rip open, rip off'

unbare   'bare'

unstrip   'strip'

unthaw   'thaw'

undecipher   'decipher'

Here the *un-* does not reverse the change of state signified; it leaves it more or less the same. Of course, *unbare* may be a calque of French *dénuder*, reflecting the conflation in Romance languages of Latin *dis-* 'un-' and *dē-* 'down, off'; and *unpick* and *unrip*, which go back to the fourteenth and sixteenth centuries respectively, may represent an older sense of *un-*, as in *unloose*. But *unthaw, unstrip,* and *undecipher* are genuine problems, and the fact that they cannot be explained away argues against trying too hard to explain away the others.

The relevant observation is that the verbs that pick up redundant *un-* have meanings of the form CAUSE TO BECOME NOT *X*, so adding *un-* to them removes exceptions to the surface generalization that verbs that begin with *un-* mean CAUSE TO BECOME NOT *X*. There is, however, no way to state this fact in the impoverished-entry theory. Either *unthaw* is not derived from anything, in which case the fact that part of its meaning is predictable from the presence of *un-* is ignored; or else *unthaw* is derived from *thaw* by a "redundant-*un-* rule" distinct from the *un-* rule above in that it introduces no change of meaning. In either case a generalization is lost.

The full-entry theory fares much better, provided one makes the assumption that *unthaw* (for instance) has come into existence by REPLAC-ING *thaw* in the lexicon (dialectically or idiolectically), rather than being added alongside it. When this happens, the independent infor-mation content of the lexicon is reduced, for the *un-* rule formulated above (the regular one, not a special rule for redundant *un-*) can now apply rootlessly, showing that the CAUSE TO BECOME NOT in the meaning is predictable from the *un-* in the morphology. Redundant derivation thus amounts to a simplification, and the full-entry theory explains why there is functional pressure in favor of it. There is of course nothing to prevent dialects that have *thaw* and dialects that have *unthaw* from subsequently merging, so that both forms exist in the lexicon.

Other examples of redundant derivation are amenable to similar explanations. For instance, German *Hinde* 'female deer' and *Prinzess* have picked up the suffix -*in* 'female', giving *Hindin* and *Prinzessin*; this change makes semantic and syntactic gender predictable from the ending. Likewise, *thus* picks up the -*ly* suffix to give *thusly*, which is more obviously a manner adverb. In dialectal English, forms like *ungodless*, *uncomfortless*, and *musicianer* have been recorded (Botkin 1931). All of these are plausibly analyzed as reductions in the inde-pendent information content of the lexicon, under the full-entry theory.

Incidentally, there is also redundant inflection, as in the child-language forms *feets* and *camed*. Redundantly inflected forms can come into general use; thus the singular-plural pair *cow:cy* gave way to *cow:kine* on the analogy of *ox:oxen*, and German *gessen* (*ge-essen*) has been remodeled as *gegessen* (Anttila 1972:92). This perhaps provides an argument for handling inflection, or at least some inflection, with a mechanism similar to that used for derivation.

2.3 Aronoff's word-formation rules

The theory of derivation proposed by Aronoff (1976) is essentially a variant of full-entry theory, though Aronoff refers to LRRs as word-formation rules (WFRs) and is interested mostly in the phonological, rather than syntactic and semantic, aspects of their operation.

The Lexicalist Hypothesis (FET or IET) entails that lexical items are formed from other lexical items, not from an abstract repertoire of morphemes. Aronoff (1976:23-7) shows that this, in turn, entails an interesting and independently motivated constraint on the phonological cycle. Chomsky and Halle (1968, chapter 2) have proposed that derived lexical items have internal constituent structure, and that phonological rules apply cyclically, first on the innermost bracketed structures, and then outward level by level. Thus, the stress rules apply to *prohibition* as follows:

(2.3-1) Underlying representation $[\ [\ \text{prōhibit}]_V \ \text{ion}]_N$

       Main Stress Rule   - 1 -

       Main Stress Rule   - 2 1   -

       Auxiliary Reduction Rule   2 3 1   -

On the first cycle, the Main Stress Rule applies only to [prōhibit], and on the second cycle it applies to the whole word. If [prōhibit] were not bracketed off, there would be no first cycle, and there would be no stress at all on the second syllable; the *h* would elide out, and *prohibition* would be pronounced [proəbišən] rather than [prohibišən]. Brame (cited by Aronoff 1976:25) notes that [proəbišən] actually occurs in the sense 'period during which alcoholic beverages were illegal in the United States'; it is plausible to assume that, since the connection with *prohibit* is weaker in this case, the internal bracketing has been lost.

Brame has proposed a constraint on such bracketing as follows (Aronoff 1976:24):

(2.3-2) A proper substring of a given phonological string can be bracketed only if (1) the bracketed substring occurs elsewhere as an independent word, and (2) the meaning of the whole is a compositional function of the meanings of the bracketed and unbracketed substrings.

(This is not a direct quote.) Thus, [ex]*ploit* is not an acceptable bracketing, since *ex* (at least in the sense intended) is not an independent word, and [fil]*ter* will not do either, since the meaning of *filter* cannot be described as a compositional function of *fill* and *ter*. Brame gives phonological motivation for this constraint.

Brame's constraint follows from the Lexicalist Hypothesis provided that each bracketing is taken as resulting from the application of one word-formation rule; that is, the phonological material carried over from the root is bracketed. Thus, [*prohibit*]*ion* is derived by WFR from *prohibit*, and [[*prohibit*]*ion*]*ist* from that. Assuming that Brame's

constraint is well motivated, then this fact constitutes a success for lexicalism--but not necessarily a failure for a transformational account of derivation, which could easily be constrained in a similar way.

As an argument for the full-entry version of lexicalism, Aronoff (1976:28-9) adduces some facts about irregular back formation. Speakers can reconstruct a missing root on the basis of its derivatives and add it to the lexicon; thus *peddle* was back-formed from *peddler* (itself a mis-analysis of *pedlar*), and *aggress* (from *aggression* et al.) is coming into use. We have already considered Jackendoff's argument that the roots of rootless derivatives such as *conflagration* are not listed in the lexicon. Aronoff's argument, which is supposed to lend stronger support to the same claim, concerns 'wrong' back formations, such as *self-destruct* (not *self-destroy*) from *self-destruction*, and *cohese* (not *cohere*), from *cohesion*. These, he says, show that *self-destruction* and *cohesion* are entities unto themselves in the speaker's mental lexicon, and not the output of rules operating on non-lexically-insertable roots, as proved by the fact that when the speaker tries to produce the root, he gets it wrong.

However, Aronoff cites no evidence that the speaker makes the mistake at the time of coining the back formation, rather than when he initially enters the derivative in his lexicon. Perhaps the speaker, in learning the word *cohesion* (for example), fails to realize that it is derived from *cohere* and instead derives it, IET-style, from a non-insertable root *cohese*, which he makes up on the spot. Thus, Aronoff's argument does not in fact discriminate between FET, IET, or even the transformational account of derivation.

The biggest difference between Aronoff's and Jackendoff's versions of FET (aside from the fact that Aronoff has much more to say about phonology) is in the handling of completely regular, completely productive derivational processes, such as the English rule that attaches *-ness* to adjectives to produce abstract nouns. There seem to be no, or practically no, accidental gaps among *-ness* derivatives, and no unpredictable changes of meaning. That is, *-ness* derivatives contain no independent information; their entire lexical entries are completely predictable. Even the information that the derivative exists amounts to nothing, since each derivative could not help but exist: there are no accidental gaps. Thus, a Jackendoffian treatment of *-ness* derivatives requires a large number of lexical entries with absolutely no independent information content. Aronoff posits that such informationless lexical entries do not exist, and that instead the *-ness* rule (and every completely productive derivational rule) is used to form the derivatives generatively, rather than functioning as a redundancy rule.

Support for this hypothesis comes from the phenomenon of BLOCKING, i.e., the fact that it is impossible to form two derivatives with

synonymous affixes from the same sense of the same root. (This of course follows from the more general principle that languages do not tolerate exact synonymy.) Aronoff points out that there is one situation in which blocking fails to apply and exact synonymy does occur: the situation in which at least one of the synonyms is derived by a completely productive process, as in the pairs *tenacity:tenaciousness*, *speciosity:speciousness*, *impecuniosity:impecuniousness*, and the like. If perfectly regular derivatives are not listed in the lexicon, but other derivatives are, then it is possible to say that languages do not tolerate exact synonymy IN THE LEXICON, and all is well. (For more on blocking see Broselow 1977.)

3. Conclusions: Lexicalism in perspective

It is now time to put all this information together and take a stab at answering the original question: how well motivated is lexicalism?

To begin with, the Lexicalist Hypothesis does not, by itself, make a clear enough claim to be tested. Rather, arguments for and against lexicalism amount to arguments for and against particular analyses constructed within a lexicalist framework. The clearest contrast between frameworks is that between full-entry lexicalism and the transformational treatment of derivation; impoverished-entry lexicalism lies somewhere in between.

One major line of argumentation is that if derivation is handled transformationally, then the needed transformations turn out to have special properties that distinguish them from all other transformations--properties which lexical rules have automatically. For instance, it is argued that derivational rules are ordered before all other transformations, are alone capable of changing grammatical categories (e.g., verb to noun), and are required to produce the bracketing needed by the phonological cycle.

Perhaps the strongest evidence along this line comes from derivational idiosyncrasy: derivational rules have to be able to produce forms with semantic content that is not predictable by rule. If derivation is to be done by transformations, then various roots have to be listed in the lexicon with special senses that are marked as positive absolute exceptions to derivational rules. For instance, there has to be a sense of *edit* that is obligatorily transformed into *edition* (as in *hardback and paperback editions*) and never appears as a verb. These exception features on special senses can control only derivational rules, and not (for example) Gerundive Nominalization. If, on the other hand, derivatives have their own lexical entries, then their idiosyncrasies

can be listed there, and the need for restricting the exception features to derivational transformations disappears along with the transformations themselves.

On the whole, the evidence for such a contrast between derivational rules and (other) transformations is reasonably strong, and thus the Lexicalist Hypothesis is well supported. Interestingly, such a contrast argues against, rather than for, the more recent proposal that all relation-changing transformations should be replaced by lexical rules, since such a move would obliterate the contrast, or put it in the wrong place. (Such a theory is proposed by Bresnan (1978), as noted above.)

The other major line of argumentation concerns the claim that lexicalist accounts of derivation are more psychologically real than transformational accounts. Besides the various plausible but untested claims about learnability and the like adduced by Jackendoff (1975), this line also includes diachronic arguments such as Jackendoff's argument from *gooseberry* ‹ *groseberry* and my argument from redundant derivation. Again, lexicalism is reasonably well supported by the evidence.

With regard to psychological claims, lexicalism proper and pan-lexicalism are closely parallel; in particular, the full-entry lexicalism of Jackendoff (1975), the extended lexicalism of Bresnan (1978, especially p. 14), and the "space grammar" of Langacker (1980) all lend themselves well to performance models in which the brain is a slow computer with a large long-term memory and uses rules to supplement, rather than replace, the storage of large amounts of information about individual lexical items. (Traditional transformational grammars, on the other hand, read as though they were optimized for fast computers with small memories, making maximal use of computation rather than look-up.) Of course, a great deal of rigorous testing is needed to give some solidity to these speculations.

NOTES

[1] I shall address only arguments that bear directly on the issue; I am therefore not concerned with giving a full picture of the content and emphasis of each paper I review.

I am indebted to Margaret M. Fleck, David Johnson, Donald Ringe, and Rulon Wells for many helpful comments and suggestions. However, they share no responsibility for my mistakes and do not necessarily agree with my conclusions.

[2] Note that *John's likelihood of winning* is not a counterexample. It is derived by raising out of a gerundive, from *the likelihood of John's winning*.

[3] Deriving *blue jeans by Calvin Klein* by ellipsis from *blue jeans designed by Calvin Klein*, or the like, does not solve the problem. To capture the fact that (1.1-7) and (1.1-8) are parallel, we would then have to derive *the destruction of the city by the enemy* from something like *the destruction of the city brought about by the enemy*, and NOT from a passivized embedded sentence.

[4] Postal in fact also makes another semantic argument based on quantifier scope. I find the argument I describe considerably the clearer of the two.

[5] An argument of this sort seems to have been advanced by Chomsky (see McCawley 1975:213, footnote), but the details of this one are my own.

[6] There are, of course, derived nominals that end in *-ing* (*John's multiple rereadings of the book, readings from classical authors, the failings of great men*), but these show all the syntactic properties of derived nominals and also show derivational idiosyncrasy (note the difference in concreteness between *readings* and *writings*), so the distinction is intact.

[7] This fact suggests that if the Lexicalist Hypothesis is to be adopted, the lexicon should use the same mechanism to relate the senses of any given lexical item and to deal with semantic idiosyncrasies between root and derivative.

[8] All the examples here are from Sadock.

[9] In fact, negative *un-* and reversative *un-* have different etymological origins. The former is cognate with German *un-* and Greek *a-* (alpha privativum), the latter with German *ent-* and Greek *anti-*.

[10]The English lexical passive rule thus displays an ergative pattern: the subject of an intransitive verb like *travel* and the object of a transitive verb like *notice* are treated in the same way. In this, it parallels the *-ta* participle rule of Sanskrit, which produces participles that mean 'having been *X*ed' when the verb is transitive and 'having *X*ed' when it is intransitive. This participle is the source of the ergative past tense in Hindi.

[11]These are Nanni's grammaticality judgements; *pattern* is presumably meant here as a sewing term. Many speakers find *an easy-to-make-looking pattern* ungrammatical; perhaps *a hard-to-fool-looking man* is better.

[12]One of the principal weaknesses of Jackendoff's exposition is the ambiguous use of the term *intuition*, which seems to refer at various points to the theoretician's feel for how things ought to be analyzed, the native speaker's intuitions of lexical relatedness and of other things, and various plausible psycholinguistic claims. I have tried to pick these apart. As far as I can tell, Jackendoff's ambiguity is relatively superficial and does not hide any basic confusions.

[13]Actually, the *-ion* rule needed here is somewhat different from (2.1-3), since *conflagration* takes neither a notional subject nor a notional object. Perhaps the NPs in the syntactic frame in the left-hand side of (2.1-3) need to be optional.

[14]Jackendoff in fact proposes to separate morphosyntactic and semantic LRRs completely (1975, section 4). This obviously will not do, for on that analysis it would not be surprising if *untie* meant 'tie again' and *retie* meant 'untie' (with all other *re-* and *un-* verbs as they are now).

REFERENCES

Anttila, Raimo. 1972. *An Introduction to Historical and Comparative Linguistics.* New York: Macmillan.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar. Linguistic Inquiry Monograph Series,* 1. Cambridge, MA.: MIT Press.

Botkin, B. A. 1931. Folk speech in the Kentucky Mountain Cycle of Percy MacKaye. *American Speech* 6.264-76.

Bresnan, Joan. 1978. A realistic transformational grammar. In Morris Halle, Joan Bresnan, and George A. Miller (eds.), MIT Bicentennial Studies, 3. Cambridge, MA.: MIT Press, pp. 1-59.

Broselow, Ellen. 1977. Language change and theories of the lexicon. *Papers from the Thirteenth Regional Meeting of the Chicago Linguistic Society,* pp. 58-66.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs and Peter Rosenbaum (eds.) *Readings in English Transformational Grammar.* Waltham, MA.: Ginn, pp. 184-221.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English.* New York: Harper & Row.

Covington, Michael A. 1980. Redundant derivation: diachronic evidence for the fully specified lexicon. Paper presented at Linguistic Society of America annual meeting.

Emonds, Joseph E. 1970. Root and structure-preserving transformations. Dissertation, MIT. Formerly distributed by the I.U. Linguistics Club.

Freidin, Robert. 1975. The analysis of passives. *Language* 51.384-405.

Hust, Joel R. 1978. Lexical redundancy rules and the unpassive construction. *Linguistic Analysis* 4.61-89.

Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, MA.: MIT Press.

_____. 1975. Morphological and semantic regularities in the lexicon. *Language* 51.639-71.

_____. 1977. $\overline{\text{X}}$ syntax: a study of phrase structure. *Linguistic Inquiry Monograph Series,* 2. Cambridge, MA.: MIT Press.

Katz, Jerrold J., and J. A. Fodor. 1963. The structure of a semantic theory. *Language* 39.170-210.

Langacker, Ronald W. 1980. Explorations in space grammar. Lecture given at the LSA Summer Institute, Albuquerque, June 23.

McCawley, James D. 1975. Review of Noam Chomsky, Studies on semantics in generative grammar. *Studies in English Linguistics* 3.209-311.

Nanni, Deborah L. 1980. On the surface syntax of constructions with easy-type adjectives. *Language* 56.568-81.

Newmeyer, Frederick J. 1971. The source of derived nominals in English. *Language* 47.786-96.

_____. 1980. *Linguistic Theory in America: the First Quarter-Century of Transformational-Generative Grammar*. New York: Academic Press.

Postal, Paul M. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge, MA.: MIT Press.

Rischel, Jørgen. 1971. Some characteristics of noun phrases in West Greenlandic. *Acta Linguistica Hafniensia* 13.213-45.

Ross, John Robert. 1973. Nouniness. In Osamu Fujimura (ed.) *Three Dimensions of Linguistic Theory*. Tokyo: TEC, pp. 137-257.

Sadock, Jerrold M. 1980. Noun incorporation in Greenlandic: a case of syntactic word formation. *Language* 56.300-319.

Wasow, Thomas. 1977. Transformations and the lexicon. In Peter W. Culicover, Thomas Wasow, and Adrian Akmajian (eds.) *Formal Syntax*. New York: Academic Press, pp. 327-60.