The Lexicalist Hypothesis: Both Wrong and Superfluous

Benjamin Bruening, University of Delaware (bruening@udel.edu)

draft, October 20, 2016; comments welcome

Abstract

The Lexicalist Hypothesis, which says that the component of grammar that produces words is distinct and strictly separate from the component that produces phrases, is both wrong and superfluous. It is wrong because (1) there are numerous instances where phrasal syntax feeds word formation; (2) there are cases where phrasal syntax can access sub-word parts; and (3) claims that word formation and phrasal syntax obey different principles are not correct. The Lexicalist Hypothesis is superfluous because where there are facts that are supposed to be accounted for by the Lexicalist Hypothesis, those facts have independent explanations. The model of grammar that we are led to is then the most parsimonious one: there is only one combinatorial component of grammar that puts together both words and phrases.

1 Introduction

The Lexicalist Hypothesis, usually attributed to Chomsky 1970, is a foundational hypothesis in numerous current approaches to morphology and syntax, including Head-Driven Phrase Structure Grammar (HPSG), Lexical Functional Grammar (LFG), Simpler Syntax (Culicover and Jackendoff 2005), various versions of the Principles and Parameters and Minimalist models, and others. The basic tenet of the Lexicalist Hypothesis is that the system of grammar that assembles words is separate from the system of grammar that assembles phrases out of words. The combinatorial system that produces words is supposed to use different principles from the system that produces phrases. Additionally, the word system is encapsulated from the phrasal system and interacts with it only in one direction, with the output of the word system providing the input to the phrasal system. This has the result that the phrasal system has no access to sub-word units, and in addition, the output of the phrasal system never forms the input to the word system.

An alternative view dispenses with the Lexicalist Hypothesis and argues that the phrasal and word formation systems are not distinct (Sadock 1980; Baker 1985; Sproat 1985; Lieber 1988, 1992; Hale and Keyser 1993; Halle and Marantz 1993; Marantz 1997; Borer 2005; Bruening 2014; among others). If this view is correct, a model of grammar does not need two separate components, but only one. This view has the virtue of simplicity: a model of grammar with only one component is simpler than one with two, and is therefore to be preferred, assuming that they are equivalent in their empirical coverage. In the face of this challenge, numerous recent publications have defended the lexicalist position, arguing that the empirical facts demand the strict separation of the word and phrasal systems (e.g., Ackema and Neeleman 2004; Williams 2007; Newmeyer 2009; Müller 2013; Müller and Wechsler 2014).

---

1This hypothesis is usually referred to as the Lexicalist Hypothesis, after Chomsky (1970). It is occasionally also referred to as the Lexical Hypothesis (e.g., Williams 2007). The more common name seems to be the Lexicalist Hypothesis, so that is the name I will use. The literature on this hypothesis is too vast to cite fully here. Important references include Jackendoff (1972), Aronoff (1976), Lapointe (1980), Bresnan (1982), Kiparsky (1982b), Simpson (1983a), Mohanan (1986), Di Sciullo and Williams (1987), and Bresnan and Mchombo (1995).

2A necessary assumption of this view is that it is possible to define the word as a linguistic unit and to unambiguously identify words. As is well-known, this assumption is false. There is no universally valid definition of a word, and even in a well-studied language like English, there are numerous cases where it is unclear what constitutes a word. See Haspelmath (2011) for recent discussion. I will put this issue aside in this paper, and show that, even if we grant the lexicalist assumption that it is possible to identify words, the lexicalist position is both wrong and superfluous.
In this paper I make two points. First, the empirical facts indicate that the Lexicalist Hypothesis is fundamentally incorrect. There are numerous instances where the output of the phrasal system feeds word formation (section 2), and there are also cases where the phrasal syntax must have access to sub-word units (section 3). These facts are incompatible with the model of grammar posited by the Lexicalist Hypothesis. Morphology and syntax also do not obey different principles, but instead behave alike (section 4). This removes one argument for the Lexicalist Hypothesis. Second, the Lexicalist Hypothesis is superfluous. Where there are facts that are supposed to be accounted for by the Lexicalist Hypothesis, those facts have independent explanations, even within theories that assume the Lexicalist Hypothesis. This is shown in section 5. Since the Lexicalist Hypothesis is both incorrect and does no work, it can and should be dispensed with. We only need a single syntactic module in the grammar, one that produces both words and phrases.

Before beginning, I should point out that different approaches that assume the Lexicalist Hypothesis vary significantly in their implementation. Throughout this paper, I try to make points that apply to all of them. This is not always possible, however, so where relevant I try to be clear about which specific approaches or claims I am criticizing, and note at that point that the criticisms may not apply to certain other approaches within the Lexicalist Hypothesis. I should also point out that there is another meaning of the word “lexicalist” as applied to theories, where it means that that theory has information encoded in lexical entries do much of the work of the grammar. Such approaches are referred to as “lexicalist” because they include very rich lexical entries. This paper has absolutely nothing to say about this meaning of the term “lexicalist.” It addresses only the claim that there are distinct word and phrase systems in the grammar.

To begin, in sections 2–4 I point out three ways in which the Lexicalist Hypothesis is mistaken in its view of grammar.

2 Error 1: Phrasal Syntax Can Feed Word Formation

According to the Lexicalist Hypothesis, interaction between the word and phrase systems is unidirectional: the output of the word formation system provides the input to the phrasal syntax, and not vice versa. In this section I show that this is incorrect: there are numerous cases where the output of the phrasal syntax can feed word formation.

2.1 Words Obviously Containing Phrases

I begin with words that quite clearly contain phrases within them. This occurs with compounds, with zero-derivation, and with overt derivational morphology. Many of these cases are well-known and have been discussed at length. All of them have been claimed to not be problematic for the Lexicalist Hypothesis, but I argue here that they are.

The most well-known case of phrases inside words involves compounding. Nominal compounds in numerous languages may include a phrase as the first member of the compound (Botha 1981; Kiparsky 1982a; Selkirk 1982; Fabb 1984; Sproat 1985; Lieber 1988; 1992; Spencer 1988; 1991 413–417). Phrasal compounding has been documented in Afrikaans (Botha 1981), Dutch (Ackema and Neeleman 2004), German (Meibauer 2007), Mandarin Chinese (Wiese 1996), Japanese (Shibatani and Kageyama 1988), and English. I illustrate the phenomenon with English:

Chomsky (1970) presents this in derivational terms: rules of the “base” create lexical items, which are plugged into deep structures in the phrasal system; deep structures undergo transformations to surface structures, with no returning to the rules of the base. According to Wasow (1977, 330) this ordering follows from the architecture of the Extended Standard Theory, “which requires that lexical rules relate items in the lexicon, while transformations must operate on phrase markers into which lexical items have been inserted.” A similar feed-forward view is presented in Kiparsky (1982b) and Mohanan (1986). In constraint-based frameworks, there is no derivational directionality, but the interaction is still only one way. For Bresnan and Mchombo (1995, 192) the morphological component is “restricted to lexical and sublexical categories,” meaning that the output of the phrasal system, phrases, is not something that the lexical system can work with. A similar view, where the word and phrase systems are entirely distinct and operate on different units, is presented in Di Sciullo and Williams (1987, 48–49). In HPSG, it is not entirely clear what prevents phrases forming the input to lexical rules; most HPSG work just seems to assume the lexicalist position. Sag (2007) stipulates that lexical rules/constraints involve the type lexeme; this stops them from including phrases.
(1)  
  a. I gave her a don’t-you-dare! look.  
  b. She baked her fiance a sweet I-love-you cake.

In most theories, compounding must be a lexical process, since it creates elements that have the distribution (and inflectional morphology) of words. It can also feed further derivation, as in snowboarder, pan-fryable, antibullfighting, Church Slavonicism, redownload. The lexicalist literature uniformly treats compounding as a lexical process (e.g., Kiparsky 1982b, Bresnan and Mchombo 1995). Compounds should therefore not be able to include phrases that are put together by the phrasal syntax.

In fact, however, the phrases that constitute the first member of these compounds must be put together by the syntax, because they have the form that the syntax requires. They can be imperatives and have the form of an imperative, as in (1a), and they can have the form of a declarative, as in (1b). They can also be wh-questions and exclamatives, as in the examples in (2). In all cases, if the first member of the compound violates rules and constraints of the phrasal syntax, the result is ill-formed (3):

(2)  
  a. She had that Don’t-you-dare! look.  
  b. She had that I’m-so-proud-of-myself look.  
  c. She had that What-the-hell-are-you-doing?! look.  
  d. She had that What-a-strange-person-you-are! look.

(3)  
  a. * She had that You-don’t-dare! look. (obligatory inversion with negative imperative)  
  b. * She had that Myself-is-so-proud-of-me look. (reflexive bad as subject)  
  c. * She had that You’re-doing-what-the-hell?! look. (obligatory fronting with wh-the-hell; Pesetsky 1987)  
  d. * She had that What-a-strange-person-are-you! look. (no inversion with exclamatives)

If these phrases were not put together by the phrasal syntax but by some other mechanism, that mechanism would have to precisely duplicate the constraints of the phrasal syntax.

It is also not the case that the phrases that appear inside compounds are “lexicalized” units, as Bresnan and Mchombo (1995) claimed. Using phrases in compounds is completely productive, as the examples above and attested usage show (see Ackema and Neeleman 2004, Lieber and Scalise 2007, Sato 2010, and especially Hohenhaus 1998). Here are a few attested examples, pulled from the web (with the punctuation used there):

(4)  
  a. What was your “I don’t get paid enough for this shit” moment?  
      (https://www.reddit.com/r/AskReddit/comments/3hiw5t/what_was_your_i_dont_get_paid_enough_for_this/)  
  b. How to end your “I don’t feel like it” syndrome  
      (http://www.prolificliving.com/i-dont-feel-like-it-syndrome/)  
  c. If there’s one thing I don’t need, it’s your “I don’t think that’s wise” attitude.  
      (http://www.gotfuturama.com/Multimedia/EpisodeSounds/2ACV02/)  
  d. Discovering the answer to your why can’t I sleep question is an important step to putting sleep troubles behind you.  
      (http://www.holistic-mindbody-healing.com/facts-about-insomnia.html)  
  e. Your “Why can’t I bait newbies?” tears are glorious.  
      (https://forums.eveonline.com/default.aspx?g=posts&m=5798278)  
  f. Overcoming the I can’t afford it sales objection.  
      (http://www.servextra.com/overcoming-the-i-cant-afford-that-sales-objection/)

Product ads also regularly use novel phrasal compounds; here are some noticed recently:

(5)  
  a. Lysol is “the prevent-mold-on-the-shower-curtain-for-up-to-seven-days spray”  
  b. on Puffs box: “open for the I-feel-better-already tissue”
c. “Growing Kids? The Yellow Pages is your oh-boy-they-need-more-shoes-and-clothes-and-we-should-start-braces-for-their-teeth-now directory.”

Wiese (1996) argues that phrasal compounds are not a problem for the Lexicalist Hypothesis because the phrase is a quotation. This view is supported by the fact that first and second person pronouns do not refer to the speaker and hearer in compounds like *an I-love-you cake*, just as they do not inside quotations (Bresnan and Mchombo 1995). This analysis is further elaborated by Pafel (2015). I will return to this idea momentarily, after illustrating other instances of phrases inside words.

Words can also be zero-derived from phrases, another instance of what is supposed to be banned by the Lexicalist Hypothesis. These phrases can include a verb plus adverb, verb plus object and particle, verb plus preposition, modal verb plus main verb, or other elements, including functional ones that do not typically participate in word formation (a how-to). Most of the output forms are nouns, but at least three are adjectives (*see-through*, *lackluster*, *as-is*). The following are some examples that include at least three different words:

(6) a. a ne’er-do-well
   b. a know-it-all
   c. a wannabe (want-to-be)
   d. a good-for-nothing
   e. a pick-me-up
   f. a hand-me-down
   g. a shoot’em-up
   h. the pushmi-pullyu (push-me-pull-you)
   i. a free-for-all
   j. a two-by-four
   k. a mother-in-law, brother-in-law, etc.; the in-laws
   l. a (fine) how-do-you-do

There are also many examples with two different words:

(7) a. a has-been
   b. an also-ran
   c. will-call (“the tickets are waiting at will-call”)
   d. the once-over
   e. a do-over
   f. a look-alike
   g. X’s say so
   h. the long ago
   i. a gimme (give-me)
   j. a know-nothing
   k. a readme
   l. a thank you
   m. a how-to
   n. the go-ahead (“give someone the go-ahead”)
   o. a walk-through, a run-through, a drive-through, follow-through

---

4 One might also want to include in this list such words as *pickpocket*, *cutthroat*, and *scofflaw*. However, these are missing elements that they would have to have as phrases (determiners or plural morphology, and a preposition in the case of *scofflaw*). In the list in the text I include only words that are well-formed phrases.
p. see-through
q. lackluster
r. as-is

A particularly large and productive class consists of verb-particle combinations:

(8) a drive-by, a layabout, a runaway, a knock-off, a callback, a rub-down, a beat-down, a smash-down, a put-down, a pushover, a cast-off, a castaway, a breakup, the brush-off, a let-down, a walk-up, a walkabout, a throwback, leftovers, . . .

These verb-particle combinations contrast with ones where the particle is initial, like input, bystander, off-putting. They therefore do not appear to have been formed by the process of compounding. Also fully productive are adjectival passives consisting of verbs plus particles (put upon, run down, crossed out, etc.), as well as adjectival passives derived from verbs plus full prepositional phrases (talked about, unasked for, unheard of, etc.). See section 2.2 on these prepositional passives. There is also a productive process for forming nouns denoting food and beverage dishes and names of products and services from phrases, primarily conjunctions of noun phrases. Some examples include two gin and tonics, a Stoli and kiwi juice (Wechsler 2008b), a surf-n-turf, an East meets West, a bed-and-breakfast, the wash-n-fold, a slip-n-slide, etc.

Verbs can also be created from phrases. Here are a few attested examples:

(9) a. You just Bonnie and Clyded my starting middies! (Archer season 3, episode 3)
   b. The outside door opened and a knot of people, all taking down umbrellas and shaking out hats, came in, were order-of-serviced by Colin, and went into the nave. (they were handed a card listing the order of service; Connie Willis, The Doomsday Book)
   c. Don’t you ‘wee, tim’rous beastie’ me. (New Yorker cartoon, cited by Clark and Gerrig 1990, 773, (8a))

See also Carnie (2000) on verbs zero-derived from phrases. Besides the adjectives listed above, Ackema and Neeleman (2004, 177, note 25) note a productive use of phrases as adjectives in certain currently popular registers:

(10) a. This is too [A last year] to wear.
    b. That music is so [A early eighties].

Many words derived from phrases can feed further derivation. For instance, knowitallism and ne’er-do-well-ism are well-attested. Clark and Gerrig (1990, 773, (8c)) cite the following example:

(11) He let it obsess him, for all the irritated now-nowing of his wife and the confusion and unease of his children. (Martin Amis, Success)

Here now-nowing is a noun derived from a verb zero-derived from a phrase. In fact, any verb derived from a phrase could similarly be turned into a noun with the suffix -ing (Bonnie-and-Clydeing, etc.).

The simplest analysis of all these cases is that they are phrases, put together by the phrasal syntax, and then converted into nouns (or adjectives, or verbs).\(^5\) The ones that are nouns are clearly nouns, since they occur with articles and/or possessors and can take plural morphology (a couple of wannabes, two has-beens, most how-tos, three brother-in-laws). The ones that are verbs take verbal morphology (Bonnie-and-Clyded). However, in every

\(^5\)In a purely syntactic approach to word-formation, one possible analysis would be to posit a null nominal head that merges with the phrase and turns it into a noun. An anonymous reviewer says that the purely syntactic approach to morphology necessarily adopts this kind of syntagmatic approach to morphology and could not adopt a paradigmatic one like that advocated by Becker (1993). It remains to be seen whether such an approach could handle all of the facts listed in Becker (1993). It is true, as the reviewer also notes, that a syntagmatic approach will need to admit empty heads/morphemes, as just suggested, but whether this is a drawback or not seems to be entirely a matter of opinion. It seems to me that empty elements are widespread in syntax and morphology and cannot be denied, so my own view is that empty elements in morphology are not a drawback at all. Of course, the question will always be what analysis achieves the best empirical coverage.
case the word has the form produced by the phrasal syntax, and not the form that ought to have been produced by,
say, compounding (e.g., *it-all-know, *so-say). Many of them also include non-selected adverbs (a do-over, an
also-ran), which in some Lexicalist approaches are supposed to be inaccessible to the word formation system (see
section 4.2). Some include functional elements that do not otherwise appear inside words (a how-to, a wannabe).
The noun how-to even includes apparent wh-movement of a wh-adverb, as does a fine how-do-you-do, which also
includes subject-auxiliary inversion. This type of zero derivation, then, like the case of phrasal compounding
above, is an instance of phrasal syntax providing the input to word formation. Again, this ordering is not supposed
to exist under the Lexicalist Hypothesis.

One could attempt to dismiss these forms as isolated exceptions. However, as can be seen from the list above,
there is a large number of them, and at least two classes are completely productive (verb-particle combinations,
and names for foods/beverages and products). Compare this with the noun-deriving affix -age, as in blockage, breakage,
cleavage, leakage, lineage, stoppage, etc. This suffix occurs with what is probably a comparable number of words
(I count 150 or so), but is far less productive (new forms are not created frequently). Yet no one views forms in
-age as isolated; they are viewed as a pattern. Zero derivation of phrases is just as common, and new forms are
created, probably more frequently (e.g., a stop-and-chat, on the TV show Curb Your Enthusiasm). Zero derivation
of words from phrases therefore constitutes a pattern of word formation that requires an account. Moreover, if the
Lexicalist Hypothesis were correct in its view of grammar, there should be no such isolated exceptions: the form
of the grammar would simply preclude their ever being created.

There are also instances of phrases combining with overt derivational morphology to derive new words. These
include do-it-yourselfer, stick-to-itiveness, unputdownable, muckupable, backupable. Some prefixes also seem
to productively attach to phrases, in particular pre-, post-, and ex:- it was pre-founding of Rome, post-digestive
disorder complications (Lieber and Scalise 2007, 11), ex-Secretary of the Interior. See Lieber and Scalise (2007,[12])
discuss the Dutch affixes -achtig and -loos, which can attach to phrases but have all the qualities of an affix. An
anonymous reviewer also notes that phrases can be suffixed with -ness in English, as in His general ok-with-
less-than-we-should-aim-for-ness makes him an undesirable candidate. The TV show Archer created the adverb
agent-in-commandingly (in the episode ‘The Papal Chase’).

Many of these cases are well-known and the general phenomenon has been discussed extensively in the litera-
ture. Despite appearances, the lexicalist literature has come to the conclusion that they are not a problem, because,
as briefly mentioned above, they are quotations (Wiese 1996). As many have noticed (e.g., Cram 1978, Bresnan
and Mchombo 1995, Postal 2004, chapter 6), quotations can include phrases from other languages, non-linguistic
sounds, gestures, and even visual symbols. These can all also appear where phrases can in word-formation (com-
ounds, zero-derivation, etc.). For instance, the following are possible utterances, where what is between angled

(12) a. One person whistled <sequence of notes 1> and another person whistled <sequence of notes 2>.
    The <sequence of notes 1>-whistler was better than the <sequence of notes 2>-whistler.
    b. [One ASL interpreter to another:] I’ll <sequence of gestures> you!

One could also derive the words harumph (‘one who makes a harumphing sound’) and vroomology (‘science of
sounds that sound like vroom’) from non-linguistic sounds, if one wished.

The claim is then that the grammar includes some mechanism for taking basically anything and using it as a
Furthermore, this derived category (the Noun, in our example) is a single, encapsulated unit, the contents of which
are inaccessible to the morphology or the syntax it is embedded within. Now, the content of a phrase within a
word must be accessed, clearly, or sentences like An others-first attitude is clearly better than a me-first attitude

---

6 A reviewer suggests that zero derivation is applying not to a phrase, but to a string, based on the reviewer’s belief that examples like a how-to do not involve a complete constituent. However, such examples are constituents, they just involve null complements. Strings that are not constituents would never be allowed: *a gimme-the, *the push-my-pull-your. This indicates that zero derivation is applying to a phrase, and not a string.
would not even make sense (nor would 12a). We also see various kinds of anaphora crossing the boundaries of the quotation:

(13) a. Charles-and-Di syndrome died when she did. (coreference)
    b. He baked me a sweet I-love-you cake, but I don’t think he really does. (VP ellipsis)
    c. The old the-dog-ate-my-homework excuse won’t work because I know you don’t have one! (one anaphora)
    d. You can’t use the termites-ate-the-walls excuse because the home inspection didn’t find any! (NP ellipsis)
    e. No bag lady likes to get a “thank God I’m not her” look from passerby. (e-type anaphora)

So, the claim must be that, within the sentence or word the phrase is embedded in, it is a single, impenetrable unit, but on some parallel plane the syntax and semantics put together and interpret the phrase. Simply to be concrete, we can use Davidson’s (1984) demonstrative theory of quotation to illustrate what this might look like. In Davidson’s theory, a quotation is some kind of demonstrative in the sentence it is embedded in. This demonstrative points to the content of the quotation. We might represent this with “THIS” and “THAT” indicating the demonstrative:

(14) An others-first attitude is clearly better than a me-first attitude.
    = A THIS attitude is clearly better than a THAT attitude. (Where THIS points to “others-first” and THAT points to “me-first.”)

The types of anaphora seen above must then be limited to the types of anaphora that work across sentences and discourse, and for the most part they are (Partee 1973, Pafel 2011, 263–265). All of the above types of anaphora work on this type of paraphrase (I use parentheticals because the pointing is supposed to take place contemporaneous with the demonstrative, but parentheticals are also generally viewed as being on a separate plane from the sentence they are embedded in):

(15) a. THIS syndrome, where THIS is “Charles and Di,” died when she did. (coreference)
    b. He baked me a sweet THIS cake, where THIS is “I love you,” but I don’t think he really does. (VP ellipsis)
    c. The old THIS excuse, where THIS is “the dog ate my homework,” won’t work because I know you don’t have one! (one anaphora)
    d. You can’t use the THIS excuse, where THIS is “termites ate the walls,” because the home inspection didn’t find any! (NP ellipsis)
    e. No bag lady likes to get a THIS look from passerby, where THIS is “thank God I’m not her.” (e-type anaphora)

So far so good, then. Where phrases can appear, non-linguistic sounds, gestures, and so on can also appear, and the contents of the phrase seem to be encapsulated from the syntax they are embedded in.

Does this make such phrases inside words not a problem for the Lexicalist Hypothesis? Words can contain phrases in the quotation approach. The grammar has an explicit mechanism for embedding a phrase inside a word, namely a mechanism of the type posited by Wiese (1996) and Pafel (2015). The very existence of such a mechanism seems, on the face of it, to be a violation of the Lexicalist Hypothesis. A proponent of the Lexicalist Hypothesis would have to claim that these cases are exceptional, and there are other cases that the Lexicalist Hypothesis would rule out. It is not clear what those cases could be, though. A grammar that contains a mechanism for creating a noun stem from anything would expect anything to be able to appear anywhere a noun stem can, and this seems to be correct. Phrases and non-linguistic sounds/gestures/symbols seem to be allowed in every case of word-formation that is productive. As far as I can tell, they are only disallowed where affixes have some particular selectional requirements, but these requirements rule out other types of stems, too. It is therefore not clear that the Lexicalist Hypothesis actually rules anything out.
The quotation view does predict that a phrase could only be possible in a position where a non-linguistic sound or gesture would also be possible, but again it is not clear that this prediction rules anything out. As just stated, all of them are allowed everywhere a noun stem is. It appears that there is absolutely no content to the Lexicalist Hypothesis’s claim that phrases cannot appear inside words in general, but are only allowed by the exceptional mechanism of quotation.

Moreover, the claim that the derived category is encapsulated appears to be nothing more than the claim that it behaves like any other morpheme. Word formation in general does not access the internal structure of the stem it operates on (but see section[3] for cases where internal structure is accessed). This is the principle of strict cyclicity in morphology [25]. For instance, the suffix -ity does not care whether its stem has a bound latinate root (as in nationality) or a free germanic root (as in frontality), nor could it. All it can see is the stem it attaches to. In both cases, that stem ends in -al, which is something -ity can attach to. If I create a word based on a sound like harumphal, then I can create harumphality. Again, the suffix -ity cannot see into the stem it attaches to. Similarly, in a compound like oxtail, the fact that the first member takes an irregular plural is invisible at the level of the entire compound. The plural is oxtails, not *oxtailen. The suffix -ness attached to childless necessarily creates an abstract mass noun, not a count noun whose plural is *childlessness(r)en. The plural of the agentine noun formed from the verb to goose is goosers, not *geeser. In every case, particularities of the internal structure of a stem are invisible at later stages of derivation and inflection. The principle of strict cyclicity in morphology means that a non-phrase is just as encapsulated as a phrase as the first member of a compound or as a morpheme in a derived word.

To illustrate this further, consider words formed from names, like Reaganomics. By turning Reagan (what would be a full noun phrase in the phrasal syntax) into a noun stem for further derivation, all the properties of the noun phrase Reagan become inaccessible. For instance, it is no longer an R-expression for Condition C. The same is true for the pronoun he in words like he-man: it is no longer a pronoun for Binding Principle B:

(16) a. Reagan changed his views over his eight years as president. In later years, he rejected Reaganomics. (he=Reagan)

b. He hates he-men. (no Condition B violation)

It is not at all clear that it would be desirable to view Reagan and he as quotations in these cases of word formation. If they are, then we might as well view every stem as a quotation in word formation, since as far as I can tell they all have the same properties.

To further make this point, consider compounds. The stems that go into a compound act exactly like the same stems in other cases of derivational morphology. In the examples below, the noun phrase squirrel(s) is an R-expression in the syntax and gives rise to a Condition C violation if it is coreferential with a commanding pronoun. The stem squirrel does not, in either a compound or a derived word:

(17) a. * They₁ detest people who chase squirrels₁. (Condition C violation)

b. They love squirrel food. (they= squirrels; no Condition C violation)

c. They like to squirrel away nuts. (they = squirrels; no Condition C violation)

If we want to claim that “encapsulation” follows from quotation, then every stem must be a quotation. Conversely, if we recognize that every stem in derivational morphology is encapsulated, then there is nothing special about phrases in morphology. Their properties just follow from the way morphology works in general.

It is also important that this same encapsulation takes place entirely within words, too. This is difficult to show in English because of its relative paucity of morphology, but the third-person singular ending -s ought to make the point. Consider the following example, from a newspaper:


Here we have a word, out-Trumped, which contains a stem derived from a name. The ending -s encodes a third-person singular subject. We know from the syntactic context that this third-person singular subject is Trump himself. Yet we see no Condition C violation within the word: it is perfectly possible for the referent of -s to be the
same as the individual denoted by the name that forms the root of the word. In other words, stems are encapsulated even within words, not just across word boundaries.

What this means is that the Lexicalist Hypothesis is superfluous, as will be shown at length in section 5. That is, the fact that phrases inside words are encapsulated follows not from the Lexicalist Hypothesis, but from the way morphology works. To say that a phrase inside a word has no internal structure is just to locate it within morphology in general. Crucially, word boundaries play no role; only morpheme boundaries do.

The Lexicalist Hypothesis also rules nothing out, as far as I can tell, once one admits a mechanism that can turn any phrase into a noun stem (or adjective, or verb). I conclude that all of these instances of phrases occurring inside words really do falsify the Lexicalist Hypothesis, contra assertions to the contrary.

2.2 Prepositional Passives

In the previous subsection I mentioned adjectival passives formed from verbs plus PPs, as in talked about, unasked for, unheard of. These appear to be instances of derivational morphology taking a phrase as input (derivation of an adjective from a V+PP phrase). As can be seen, they can feed further derivation, for instance prefixation with un-. This seems on the face of it to be contrary to the Lexicalist Hypothesis. The lexicalist literature has again concluded that they are not a problem, however. The typical account is to say that the V and the P merge together into some kind of complex predicate (e.g., [Wasow 1977] Bresnan 1982, 51, Levin and Rappaport 1986). The object of the P is then the object of the complex V-P, and is on a par with the object of any V. It becomes externalized in both verbal and adjectival passives.

Bruening (2014, 387–388) shows that this analysis is untenable. The subject of a typical verbal or adjectival passive can have a resultative secondary predicate predicated of it, since it corresponds to the direct object of the active verb (e.g., Simpson 1983b, Levin and Rappaport Hovav 1995, Embick 2004):

(19) Verbal Passives
   a. The fields are being trampled flat.
   b. The covers are being shaken smooth.
   c. The cockroaches are being stomped dead.

I also question whether phrases in compounds are always truly quotations, and whether they are always fully encapsulated. [Pafel (2015)] distinguishes quotative phrasal compounds from non-quotative ones. The latter include examples like a slept-all-day look, but Pafel does not say how to distinguish a quotative compound from a non-quotative one. [Ackema and Neeleman (2002)] also argue that not all phrases inside compounds and words are quotations, but again they do not give criteria for distinguishing them. These authors are expressing an intuition, which I share, that not all phrasal compounds are the same. Quotations in general are not all the same, either. [Cappelen and Lepore (2007)] point out that mixed quotations (ones that are integrated into a sentence; see also Clark and Gerrig 1990) can involve anaphor binding across the boundary of the quotation:

(i) Mary said that Bill “loves himself” and he does. (Cappelen and Lepore 2007, 146, (11.10))

I find that this is also possible in phrasal compounds, if the phrase is not a complete sentence:

(ii) a. Sally has that about-to-talk-to-herself look.
    b. Max has that about-to-talk-to-himself look.
(iii) a. It has a curled-up-on-itself look.
    b. He has a drank-himself-to-death look.

I also find that ‘respective’ modification is possible (cf. Davidson 1984, citing J.R. Ross):

(iv) a. Her “girls rule, boys drool” taunt applies quite well to Hillary and Donald, respectively.
    b. Old Superman and New Superman are in a cool-to-tool relationship, respectively.
    c. Tex and the Marshall are in a quick-and-the-dead situation, respectively.

Most accounts of ‘respective’ modification rely on mechanisms of sentential syntax; see [Kubota and Levine (2016)] and references there. I will leave a complete exploration of these issues to future work. What is important here is that the Lexicalist Hypothesis is incompatible with the existence of phrases inside words, and is not useful in explaining their properties when they do occur there.
Adjectival Passives

a. The metal looks hammered flat.
b. The bed covers appear shaken smooth.
c. The fields have a trampled-flat look.

If prepositional passives did involve some kind of complex predicate formation, such that the object of the P becomes equivalent to the direct object of the (complex) V, we would expect resultatives to be grammatical with them, as well. They are not. A resultative interpretation is not possible (in some cases a depictive interpretation is):

Verbal Passives

a. The fields are being marched through flat. (*resultative meaning)
b. The covers are being slept on messy. (*resultative meaning)
c. The cockroaches are being stomped on dead. (*resultative meaning)

Adjectival Passives (Bruening 2014 387–388, (67))

a. The bed looks slept in smooth. (*resultative meaning, where as a result of the sleeping, the bed becomes smooth)
cf. The bed looks (frantically) shaken smooth.
b. a swept-under/tripped-over (*bumpy) rug
cf. a kicked-bumpy rug
c. * The fields had a marched-through-flat look.
cf. a trampled-flat look
d. Her bag always had a rummaged around in (*messy) appearance.
cf. a ransacked-messy appearance

In fact, subjects of prepositional passives pattern with objects of prepositions, which also do not permit resultative secondary predicates. This indicates that there is no process of complex predicate formation, and the subject of a prepositional passive does not correspond to the object of a transitive verb.

Verbal prepositional passives can even include more phrasal material, as Postal (1986) and Baltin and Postal (1996) showed:

(23) a. The bridge was flown (both) over and under. (Postal 1986 243, note 14)
b. The bridge was flown over and then, but only then, under. (Postal 1986 243, note 14)
c. Fascism was fought for by Goebbels and (then) against by De Gaulle. (Baltin and Postal 1996 30, (9d)).

In the examples in (23), a complex predicate would have to be formed out of a verb and a conjunction of prepositions, but these conjunctions are clearly phrasal.

‘Right’-type modifiers are also possible (these are based on attested examples):

(24) a. . . . no doubt that YouTube app was a trap that was walked right into!
b. I just felt that those were big details that were flown right over.
c. I pointed out that our interest was not in this program, but in doing the best possible thing for U. S. Steel, but this was glossed right over.

One could try to claim that some of these are cases of word-part ellipsis (see section 3.1 below). This will not help, however. For instance, Postal’s example with both in (23a) would have to be analyzed as follows, with elided material struck through:

(25) a. The bridge was flown (both) over and under. (Postal 1986 243, note 14)
b. = [flown both over] and [flown both under]
This analysis would still require a phrasal input to a word, and it would require both to appear in a position where it is not actually allowed (pronouncing the above string is ungrammatical; both has to scope over the conjunction). Moreover, clear instances of word-part ellipsis do not permit this: This product was hand-(both)-made and -packed (based on Chaves 2008 (4b); see section 3.1).

I conclude from this that prepositional passives, both verbal and adjectival, must be formed from a phrasal input. They are formed from a V-PP phrasal combination (and the PP can consist of conjoined PPs and can contain modifiers). This is not consistent with the Lexicalist Hypothesis. In all lexicalist accounts, adjectival passives are formed lexically (e.g., Wasow 1977, Levin and Rappaport 1986, Horvath and Siloni 2008, Meltzer-Asscher 2011). In some lexicalist accounts, verbal passives are as well (e.g., Lieber 1980, Bresnan 1982, 1995, Levin and Rappaport 1986, Müller 2006). In a lexical account, it should be impossible to form a passive from a V-PP phrase. But this is exactly what seems to be going on here.

2.3 Resultatives

Word formation from a phrasal input also occurs with resultatives and caused motion constructions. As is well known, objects can be added to verbs that do not select objects in certain constructions, for instance in resultatives:

(26) a. They fished the pond empty.
    b. * They fished the pond.
(27) a. She danced her shoes bloody.
    b. * She danced her shoes.

As Müller (2006) shows, these resultatives quite regularly participate in adjectival passive formation and nominalization in German. Corresponding to fish the pond empty is the nominalization Leerfischung, ‘empty fishing’ (Müller 2006, 868). Corresponding to dance her shoes bloody is the adjective in (28):

(28) (Müller and Wechsler 2014, 32, (33))
    a. Er tanzt die Schuhe blutig / in Stücke.
    b. die in Stücke / blutig getanzten Schuhe
    c. * die getanzten Schuhe

As (28c) shows, the resultative is crucial to forming the adjectival passive, since Schuhe is not an argument of the verb without it. This again appears to be word formation from a phrasal source.

Müller’s own analysis is a lexical one. Consider an example of a nominalization like the following:

(29) die Leerfischung der Nordsee ‘the empty-fishing of the North Sea’ (Müller 2006, 868)

The way the lexical analysis works in Müller (2006) is that a lexical rule applies to the stem fisch- (which appears in the verb fischen, ‘to fish’). This lexical rule creates a new stem fisch2 which obligatorily takes an NP and an AP argument (in addition to a subject). Semantically, the NP object is not the logical object of the stem, but is instead interpreted as the subject of the AP. This permits the stem fisch2 to appear in a verbal frame like [die Nordsee leer fischen], ‘the North.Sea empty to.fish’.

Next, a nominalization rule may apply, creating Fischung from fisch2. Fischung inherits the arguments of the stem, so it can now appear with the adjective leer ‘empty’ and the NP ‘the North Sea’ in (29). Hence, nominalization in this analysis is not nominalization of a phrase, it is nominalization of a stem, which inherits the arguments of the input stem and licenses them in the phrasal syntax.

There is an empirical problem with this sort of approach, pointed out by Williams (2011). The problem is that arguments of nominalizations are never obligatory. This means that Fischung by itself ought to be able to mean
‘fish such that X enters state Y as a result’. However, according to Müller (2006), in fact *Fischung* can only refer to a plank on a boat. Müller concludes from this fact that *Leerfischung* must be a nominalization of *leer fisch-* (p. 869). His analysis does not actually capture this, however, because it applies the rule only to the stem *fisch*- and not to the phrase *leer fisch-*.

Of course, the lexicalist analysis could stipulate that for some reason the AP argument of the stem *fisch-* is obligatory even when it is nominalized. This would just be a stipulation, however. If we instead view the nominalization as derived from a phrasal source, we explain why the resultative AP is obligatory: the only way to form a nominalization with a resultative is to start from a phrase.

While this sort of case is not decisive, it is suggestive, and a phrasal analysis appears to have advantages over a lexical one in explaining the properties of a derived word.

2.4 Summary

This section has illustrated several cases of phrasal syntax feeding word formation, in clear violation of the Lexicalist Hypothesis. Other examples have also been presented in the literature. For instance, Pulleyblank and Akinlabi (1988) illustrate a case of phrasal syntax feeding word formation in Yoruba; Goddard (1988) argues for the necessity of phrasal syntax preceding morphology in the Algonquian language Fox; Rainer and Varela (1992) list cases of prefixes attaching to phrases in Spanish; Subramanian (1988) describes phrasal syntax feeding derivation in Tamil; Lefebvre and Muysken (1988) present such a case in Quechua; Lieber and Scalise (2007, 6–8) discuss another in Italian; and Ackema and Neeleman (2004) present several cases in both English and Dutch.

Part of the original motivation for the Lexicalist Hypothesis was the apparent lack of phrasal processes feeding word formation. Since it was proposed, however, more and more instances of phrasal processes feeding word formation have come to light. The response has often been to reassign those phrasal processes to the word formation component, as in the example of resultatives in section 2.3. What this history shows, however, is that the original motivation was incorrect. Phrasal processes can and frequently do feed word formation.

3 Error 2: Phrasal Syntax Has Access to Sub-Word Units

The Lexicalist Hypothesis is also incorrect in its assertion that the phrasal syntax has no access to sub-word units. In fact, several processes can target both phrases and sub-word units.

Raising to object is also attested, but it is much rarer, and native speakers seem less willing to accept the examples that are attested. Regardless, the point is that a re-evaluation of Chomsky’s original arguments for the Lexicalist Hypothesis reveals that many of them lose all force given later lexicalist developments (and further empirical inquiry).

In most lexicalist approaches, this follows because the atoms of the phrasal syntax are words. See the works cited above, especially Di Sciullo and Williams (1987). In other formulations, Lapointe (1980, 8) has a principle stating that “no syntactic rule can refer to elements of morphological structure”; Selkirk (1982–70) has a Word Structure Autonomy Condition that bans syntactic operations from involving categories of both word and sentence structure.
3.1 Coordination/Ellipsis of Word Parts

It has long been noted that coordination of word parts is widespread (Nespor 1985, Booij 1985):

(30) a. infra e ultrasuoni (Italian, ‘infra- and ultra-sounds’; Nespor 1985, 201)
b. Freund- oder Feindschaft (German, ‘friendship or hostility’; Booij 1985, 152)
c. Pre- and post-revolutionary France were very different from each other. (Chaves 2008, 264, (10))

Chaves (2008) argues strongly that this phenomenon is actually not coordination of sub-word parts, but co-

ordination of phrases accompanied by word-part ellipsis. Among the arguments is the fact that plural agreement

is possible in (30c), as is antecedence of the reciprocal each other. The subject in (30c) must actually be Pre-

revolutionary France and post-revolutionary France, with deletion of repeated material in the first conjunct.

Regardless, ellipsis is generally regarded as a phrasal process. Bresnan and Mchombo (1995) include it in their

list of processes that cannot target sub-word units (see section 5.2). Obviously, it can. One response is to view this

type of ellipsis as different from phrasal syntax ellipsis; for instance, Chaves (2014) treats it as a process that targets

linearized strings, not syntactic phrases. Note, however, that this ellipsis process cannot break up morphemes that

are not easily segmentable, even when the same phonological string can be stranded in a different morphological

context:

(31) a. pro-choice and -gun control (Chaves 2008, 263, (6e))
b. * (both) pro-gressive and -fessional
c. * Because he is pro-fessional and -management, he is a valuable member of our team.

(32) a. You can pre- or re-mix it.
b. * They produce cranber- and dai-ry products.

(33) a. Bing and Sydney Crosby (are not related).
b. * The room was full of bing- and bon-go players.

(34) a. red- and black-banded moths
b. * red- and ind-olent

(35) a. bi- and a-sexual
b. * the Bi- and A-cre Lane Campaigns (the Biker Lane Campaign and the Acre Lane Campaign)

(36) a. bi- and a-sexual
b. birth- and adopted sons
c. * bi- and ma-son paraphernalia (bison paraphernalia and mason paraphernalia)

In (36), for instance, we see that the strings bi- and a- can stand alone, as can the string son. But (36c) does not

work, because in these particular words bi-, ma- (similar in sound to a-), and son are not distinct morphemes. This

means that this process of ellipsis is not simply operating on phonological or prosodic strings, it must have access
to morphological structure. It does seem to be true that this type of ellipsis is sensitive to prosody (Booij 1985), but
it is also true that the constraints holding of it cannot be stated solely in terms of prosody, as just demonstrated (see
also Chaves 2008, 2014). Rather, this process of ellipsis makes crucial reference to morphology. This means that
it has access to sub-word parts.

Chaves (2008, 2014) also argues that the ellipsis process at work here is the same one that we see operating on

phrasal units in right node raising and coordinate ellipsis. Corresponding to deletion of the left member of the

second coordinate in word-part ellipsis is what is sometimes regarded as non-constituent coordination:

(37) a. [half-brothers] and [half-sisters]
b. Mary [caught a fish on Monday with a fly rod] and [caught a fish on Tuesday with a spear]. (Dowty 1988, (62)).

And corresponding to deletion of the right member of the first conjunct is right node raising:
As these examples show, these two ellipsis processes can target units larger than words, namely, phrases. At the same time, however, this process must also have access to sub-word morphemes. That is, it can see both units larger than words and units smaller than words. This is a clear violation of the Lexicalist Hypothesis: the word formation system is supposed to be encapsulated from everything that deals in units larger than the word. There should be no process that can target both phrases and morphemes. Yet this type of ellipsis does exactly that.

Another example of this situation can be found in Japanese. In Japanese, it is possible for a causative morpheme to embed a disjunction of verb phrases, as Kuroda (2003) shows:

(39) Hanako ga Masao ni uti o soozisuru ka heya-dai o haraw-aseru koto ni sita
Hanako Nom Masao Dat house Acc clean or room-rent Acc pay-Cause that to do
‘Hanako decided to make Masao clean the house or pay room rent.’ (Kuroda 2003, 455, (16))

In this example, the causative morpheme -(s)ase embeds the disjunction ‘clean the house or pay room rent’. This is clearly a syntactic phrase. At the same time, the causative morpheme has all properties of an affix (see Manning, Sag, and Iida 1999, Cipollone 2001). Either this is true disjunction embedded beneath a bound morpheme and is an example of phrasal syntax feeding word formation, or it is also ellipsis and is another example of a process that can target both phrases and sub-word units.

3.2 Focus

It is also well-known that focus can target sub-word units (Selkirk 1984, 271, Wennerstrom 1993, Artstein 2004):

(40) a. That poet is from the POST-colonial era, not the PRE-colonial one.
    b. That type of action indicates an A-moral viewpoint, not an IM-moral one.
    c. That individual is TRANS-sexual, not BI-sexual.
    d. Sex-IST, not sex-Y!

Since focus can also target phrases, this is again an example of a process that makes no distinction between phrasal and sub-word units (Artstein 2004). It will not do to say that focus operates over some representation that is orthogonal to syntactic or morphological structure, because focus is sensitive to phrase structure and to constituency. This is quite clear with scope, association with focus, and focus projection. I will illustrate with focus projection. In this phenomenon, focus stress occurs on one particular phonological unit (for instance, a syllable), but what is interpreted as being in focus can project to dominating nodes in the syntax (see, e.g., Gussenhoven 1984, Selkirk 1984). For instance, focus stress on a syllable of the direct object of a ditransitive verb can be interpreted as just the direct object being in focus, or the whole VP or the whole IP:

(41) a. The teacher gave me a PLANner yesterday, not a painting.
    b. The teacher gave me a PLANner yesterday, she didn’t send me to detention.
    c. The teacher gave me a PLANner yesterday is what happened.

It is not possible to interpret just any string of elements that includes the stressed syllable as focused:

Ackema (2014) attempts to defuse the problem Japanese causatives pose for the Lexicalist Hypothesis by claiming that the Japanese causative morpheme is actually a free form, not an affix. He claims that something can be a free form syntactically but a bound form phonologically. This appears to be the least parsimonious theory possible: there is a component of word formation for some types of derivation; a phrasal syntax for putting together phrases out of free forms; and a phonological component that can create phonological words out of free forms, mimicking the morphological component. In the absence of compelling evidence, any theory with one component, or even two, should be preferred to one with three. If one believes there are good reasons to admit phonological word formation after syntax, as in Ackema’s analysis, then one should explore the option of doing all word formation this way, getting rid of the morphological component altogether.
4.1 Strict Locality of Affixes Versus Phrases

Williams (2007) points to a putative difference between affixal *self- and phrasal *self* as a motivation for the Lexicalist Hypothesis. Phrasal *self* according to Williams may take a long-distance antecedent, so that in the following example *himself* does not need to relate the internal argument of *destruction* to the external argument of *destruction*; instead the destroyer can be unspecified:

(43) John told stories about the destruction of himself. (Williams 2007, 354, (2))

In contrast, *self-destruction* strictly requires coreference between the destroyed and the destroyer. An example like (44a) cannot instead relate the destroyed with the teller of stories. This means that, in contrast with (43), (44a) is unambiguous, and must mean (44b), not (44c):

(44) (Williams 2007, 354, (3))
   a. John told self-destruction stories. (unambiguous)
   b. John told stories about one’s destruction of oneself.
   c. John told stories about the destruction of himself.

According to Williams, such facts show that “the lexical system has no delayed resolution, but the phrasal system does” (Williams 2007, 355).

Williams is incorrect about both morphology and syntax here. In (43), *himself* combines strictly with *destruction* to fill the internal argument role of *destruction*. The NP *himself* needs an antecedent, but, as is well-known, inside NPs reflexives are free to find an antecedent based on perspective or other principles (Pollard and Sag 1992, Reinhart and Reuland 1993). If *himself* were to occur as the object of the verb *destroy*, it would strictly require covaluation between it and another argument of the verb. That is, phrasal *himself*, like affixal *self-*, generally does not allow delayed resolution. It is only in contexts like within certain kinds of NPs that phrasal *himself* can take a long-distance or even discourse antecedent.

Such a discourse antecedent is also available for some instances of affixal *self-*, contra Williams:
(45)  a. Her behavior is self-serving. (the behavior does not serve itself, it serves her)
   b. Most mission statements are self-aggrandizing. (the statements do not aggrandize themselves, they aggrandize the stater)
   c. Some of the funniest remarks are self-deprecating. (the remarks do not deprecate themselves, they deprecate the remarker)

In these examples, if self-serving (for example) were to be analyzed as $x$ serves $x$, then it is not clear how self-serving could be predicated of her behavior. All of its arguments have already been saturated, and there is no role for her behavior to fill. In other examples, the NP that is the subject of a predicate adjective with -ing must be the external argument of the related verb:

(46)  a. Her robots are self-replicating. (only: the robots replicate themselves, they cannot replicate her)
   b. Those stories are self-destroying. (only: the stories destroy themselves, they cannot destroy the storyteller)
   c. These rumors are self-sustaining. (only: the rumors sustains themselves, they cannot sustain the rumorer)
   d. These tales are self-educating. (only: the tales educate themselves, they cannot educate the teller)

In (45), then, the analysis must be that her behavior is the external argument (her behavior serves $x$), and we have delayed resolution, since self- does not relate two arguments of the stem it attaches to, instead it relates the internal argument to a non-argument. The examples in (45) are exactly analogous to Williams’s example (43) in the phrasal system. As can be seen, morphology and syntax behave exactly the same: in general there is no delayed resolution (46), but in some circumstances we see delayed resolution in both morphology and syntax. (At this point I do not have any account of what those circumstances are with affixal self-.)

Examples of delayed resolution can also be found in other word-formation processes. One example occurs in the Bantu languages discussed by Bresnan and Moshi (1990). Bantu languages are famous for their valence-changing morphology. An applicative affix (“Appl”) can add an argument, for instance a benefactive. A reciprocal affix (“Recip”) can reduce valence by unifying two arguments into one, interpreted as acting in a reciprocal manner. When these two combine in the order verb-Appl-Recip in some Bantu languages, like Chichewa, the Recip can only combine the external argument (or agent) and the argument added by the applicative affix:

(47)  Chichewa (Baker 1988a)

Ana a-na-men-y-er-an-a zigawenga.
children SM-Pres-hit-Appl-Recip-Asp ruffians
‘The children are beating the ruffians for each other.’

*‘The children are beating each other for the ruffians.’

This fits Williams’s characterization of “no delayed resolution”: the reciprocal morpheme can only combine the most local arguments, it cannot look past the applicable affix and act on an argument of the verb stem that is not local to it.

However, in some other languages, like Kichaga, this is exactly what happens:

(48)  Kichaga (Bresnan and Moshi 1990)

Wá-cháká wá-i-w’ágh-i-án-à màngi.
2-Chaga 2SM-Pres-kill-Appl-Recip-FV 1.chief
‘The Chagas are killing each other for the chief.’

In (48), the theme of the verb stem ‘kill’ and the agent added by the reciprocal morpheme are interpreted as reciprocal, skipping the benefactive argument added by the applicable morpheme in between. This is delayed resolution.

---

12 Bantu Abbreviations: Asp = Aspect; FV = Final Vowel; Pres = Present tense; SM = Subject Marker; numeral = noun class.
resolution: the argument of the verb stem is not saturated immediately, but is only combined with the agent by the reciprocal morpheme after another morpheme has been added in between. (For an analysis, see Bruening 2006.)

It is true that in most cases, affixes are strictly local in their effects. However, this is true of most cases of the phrasal syntax, too. It is not true that morphology and syntax are radically different in this respect. In fact, they seem to behave exactly the same, as a theory with only one grammatical component would predict.

4.2 Differences in How Adjuncts are Treated

Williams (2007) also points to a difference between affixal re- and phrasal again to illustrate what he considers a fundamental difference between word formation and phrasal syntax. Phrasal again may include adjuncts in its scope, but re- may not:

(49) (Williams 2007, 355, (4))
   a. John re-washed the dishes on Tuesday. (not ambiguous)
   b. John again washed the dishes on Tuesday. (ambiguous as to whether it includes on Tuesday in its scope)

According to Williams, a prefix “can have scope only over the arguments of the item it adjoins to in the word system” and not the adjuncts “because the arguments of a lexical item are represented on the item itself in some way, but adjuncts are not” (Williams 2007, 355). In contrast, an item of the phrasal syntax like again attaches to an entire phrase, which can include adjuncts.

Williams’s claim about affixes is not correct, as has already been shown. The passive morpheme can affect an NP that is not an argument of the verb (section 2.2). Nominalization in German can include a resultative, also not a selected argument of the verb (section 2.3). In Japanese, the derivational suffix - sugiro takes scope over an entire VP, including adverbs (Sadock 1991, 125). Prefixes can attach to phrases that include adjuncts in Spanish (Rainer and Varela 1992). In numerous languages, a causative morpheme can take scope over an adjunct that is interpreted as modifying the caused event. I illustrate with Venda:

(50) Muuhambadzi o-reng-is-a Katonga modoro nga dzangalelo.
    salesman 3Sg.Past-buy-Caus-FV Katonga car with enthusiasm
    ‘The salesman, eagerly, made Katonga buy the car.’ or
    ‘The salesman made Katonga [buy the car eagerly].’ (Pylkkänen 2008, 83, (8))

Other languages that permit this with morphological causatives include at least Japanese (Shibatani 1990, 313–315; Harley 2008 and references there), Bemba (Givón 1976), Luganda (Pylkkänen 2008, 119), and Finnish (Pylkkänen 2008, 116). Note that the ability of an adverb to take sub-lexical scope is not necessarily an argument against the Lexicalist Hypothesis. There are analyses of sublexical scope within lexicalist approaches (e.g., the analysis of again in Dowty 1979). However, proponents of these analyses acknowledge that affixes can take scope over adjuncts, contra Williams.

Another counterexample from English involves affixal self-, discussed in the previous subsection. The most general use of this morpheme relates a logical external argument and a logical internal argument, as follows:

(51) a. self-activated = x activates x
    b. self-supporting = x supports x
    c. self-educated = x educates x

However, as noted by Bruening (2014, 418, note 32), affixal self- can also relate an external argument and an oblique:

(52) a. self-absorbed = x is absorbed in x
    b. self-centered = x is centered on x
    c. self-involved = x is involved in x
d. self-concerned = x is concerned with x
e. self-reliant = x relies on x

(53) a. a self-addressed envelope = x addresses an envelope to x
b. a self-assigned task = x assigns a task to x
c. “Mediocrity is self-inflicted, but genius is self-bestowed.” = x inflicts mediocrity upon x, x bestows genius upon x

Most of the above examples could be considered selected oblique arguments of the base verb, but the following examples are not amenable to such an analysis:

(54) a. self-acquired property = x acquires property for x
b. a self-arranged tour = x arranges a tour for x
c. the restaurant’s self-ordering kiosk = x orders for x
d. self-pay patients = x pays for x
e. self-prepared tax returns = x prepares tax returns for x
f. self-entered measurements = x enters measurements for x
g. a self-reported history of drug abuse = x reports a history of drug abuse about/on/concerning x

A for phrase can be added to any verb and is a canonical adjunct, while obliques of the type in (54g) can be added to any verb of communication. These examples look like self-relating an argument to an adjunct.

There are also cases where affixal self- seems to correspond to an emphatic reflexive:

(55) a. self-evident = x is evident by x’s self/on x’s own
b. the self-directed study of a language = x directs the study of a language by x’s self/on x’s own
c. self-study = x studies x or x studies something by x’s self/on x’s own
d. self-originating motion/impulses = x originates by x’s self/on x’s own
e. self-arising wisdom = x arises by x’s self/on x’s own
f. self-rising flour = x rises by x’s self/on x’s own

Arise and rise are unaccusative verbs and only have one argument; self- must necessarily involve an adjunct with such verbs. Emphatic reflexives are certainly adjuncts: no verb selects one as an argument. In all of these examples, then, affixal self- is able to operate on an adjunct, or at least add one. See also [Bruening (2014, 417–418), analyzing other cases of affixal self- as operating on a raising-to-object like structure where the NP involved is not a selected argument of the verb that self- attaches to.

Williams (2007) therefore seems to have arrived at an erroneous conclusion by looking at the wrong affix. The prefix that he looks at, re-, is in fact more constrained than Williams characterizes it. Unlike self-, it cannot even include selected arguments of the stem it attaches to if they have the form of an oblique rather than a direct object [Carlson and Roeper 1980, Horn 1980, Wechsler 1990]:

(56) a. * John reput the book on the shelf. (Wechsler 1990, 9, (18b))
b. John restocked the shelf.

(57) (Wechsler 1990, 12, (30))
a. * John reclimbed over the fence.
b. John reclimbed the fence.

(58) a. * Jimmy rerelied on James.

13It is unlikely that the examples I have paraphrased with for are actually based on indirect objects. Not all of these examples permit indirect objects (pay and enter), and the semantics of a double object construction is for many speakers obligatorily caused possession. This is not the right semantics for most of these examples. (It is also not clear that indirect objects are selected arguments of the verb.)
b. * Tammy is reconcerned with her son.

This constraint does not follow from Williams’s general view of word-formation processes; it seems to be idiosyncratic to re-. This particular prefix therefore cannot be used as evidence for a weaker constraint.

At the same time, the prefix re- can include a particle in its scope (Farrell 2005, Larsen 2014 375–376):

(59) (Farrell 2005 102, (14))
   a. The plumber will be out on Monday to re-hook up the washer and dryer.
   b. Have to go re-tuck in my kids.
   c. And three days later, he was re-sworn in as governor.

While it is not entirely clear what the relation between a verb and a particle is, the particle is definitely not the direct object of the verb. Note that self- does not appear to be compatible with particles at all (*self-hooking-up washers, *a self-sworn-in governor), meaning that in this respect re- is less constrained than self-.

What this comparison between re- and self- indicates is that individual affixes must be analyzed in detail, just like individual items in the phrasal syntax. There is no general principle like the one posited by Williams; instead different affixes exhibit different behavior. Some affixes can include non-arguments, while others may not and may even be more constrained. Any analysis will have to explain the behavior of individual affixes. Analyses of many types of affixes have been proposed within purely syntactic approaches: on different causative morphemes, see Pylkkänen 2008 on a syntactic treatment of re-, see Marantz 2009 Bruening (2014, 418) suggests an outline of an analysis of self-, at least as it combines with adjectival passives. Simply assigning affixation to a lexical component does not help us to understand the different properties of different affixes. Most importantly, there is no one principle governing affixes and a different one that governs phrases.

4.3 Headedness

Williams (2007) lists two other differences that he claims hold between the word system and the phrasal system. The most important one, the claim that the word system provides input objects to the phrasal system and not vice versa, has already been shown to be false (section 2). The other is the claim that the word system and the phrasal system obey different principles, so that, for instance, the word system is head-final in English, but the phrasal system is head-initial. In fact, the driving intuition behind the syntactic approach to word formation is that this is false: principles of word formation are ones familiar from phrasal syntax (Baker 1985, Hale and Keyser 1993, among numerous others). As for head directionality, it is rather superficial, and there are numerous counterexamples in both directions in English: words can be head-initial, like verbs formed with en- (e.g., enrage, enfeeble; Lieber 1988, 214), and compounds can have equal weight for their two parts (e.g., bittersweet, deaf-mute; Lieber 1988, 218). In the other direction, phrases can be head-final, like counterexamples notwithstanding and two years ago. This difference, such as it is, is simply not significant.

4.4 Idiosyncrasy

One of Chomsky’s (1970) original arguments for a lexical treatment of nominalizations, repeated in Newmeyer (2009) is that they are not completely productive or semantically regular. That is, they show a great deal of idiosyncrasy. The following list of sample semantic irregularities among derived forms (not just nominalizations) is from Newmeyer:

(60) (Newmeyer 2009 94, (7))
   a. profess (‘declare openly’)—professor (‘university teacher’)—profession (‘career’)
b. ignore (‘pay no attention to’)—ignorance (‘lack of knowledge’)—ignoramus (‘very stupid person’)
c. person (‘human individual’)—personal (‘private’)—personable (‘friendly’)—personality (‘character’)—personalize (‘tailor to the individual’)—impersonate (‘pass oneself off as’)
d. social (‘pertaining to society’; ‘interactive with others’)—socialist (‘follower of a particular political doctrine’)—socialite (‘member of high society’)

There are also nominalizations that have no corresponding verb that they could be derived from, for instance motion (*mote), tuition (*tuit). The same is true of other derived forms (social, *soci).

An often repeated view is that lexical processes and processes of the phrasal syntax differ in productivity, and this is the basis for Chomsky’s argument. The claim is that operations of the phrasal syntax are completely productive and semantically transparent, while lexical processes are not regular, often not productive, and are frequently idiosyncratic.

Since Chomsky (1970), however, this putative difference has been shown over and over to be false. Processes of word formation can be productive and semantically transparent (e.g., Di Sciullo and Williams 1987). On the other side, irregularity is not the exclusive province of the word formation system generally. There are phrasal idioms, like kick the bucket and the shit hit the fan; there are particle-verb combinations that are interpreted idiosyncratically and are not completely productive (throw up, chew out, put up with; see Jackendoff 2002); there are numerous fixed phrases (all of a sudden, never mind) and phrasal collocations (concerted effort); there are obscure limitations on different types of A-bar movement (Sag 2010). Here again there is no difference between the word system and the phrase system, and no reason to treat them differently. In fact, if we want a uniform account of idiosyncrasy, then we have to treat them the same, otherwise we will have to have two different ways of deriving idiosyncrasy, one for the word system and the other for the phrasal system.

Moreover, simply equating idiosyncrasy with listedness does not help to understand linguistic phenomena. As an example, Reinhart and Siloni (2005) discuss derived reflexive verbs in various languages, and propose that in some languages, reflexive verbs are derived in the lexicon, while in others, they are derived in the syntax. In the syntax languages, reflexive verbs are completely productive, while in the lexicon languages, reflexive verbs are limited to a small, apparently listed, set, typically verbs of grooming (dress, shave, wash, etc.). However, simply saying that these are listed in the lexicon does not explain why they are limited and in what way. Anything at all can be listed, and a listed set can be of arbitrarily large size. Why are reflexive verbs limited to a small set of particular verbs? Why dress and shave and not tie up and choke? Why not list every single verb in the lexicon, giving the appearance of complete productivity? In other words, appealing to listedness explains nothing by itself.

Newmeyer (2009, 105) does ask a pertinent question: In purely syntactic theories of word formation, where is it recorded which affixes particular roots can combine with? That is, where is it stated that the root destroy (or destruct or whatever its base form is) combines with -ion, while grow combines with -th and criticize combines with -ization? That is, where is it stated that the root destroy (or destruct or whatever its base form is) combines with -ion, while grow combines with -th and criticize combines with -ization?

This is an important question, but it is a question to be answered, not an argument. Any theory has to answer this question, whether it assumes the Lexicalist Hypothesis or not. The Lexicalist Hypothesis does not make giving an answer any easier, since in a lexicalist theory, root and affix combinations are also done via rule, just as in a purely syntactic theory. The rules just have a different name. Particular roots still have to be specified as undergoing some rules and not others. The same has to be done in a syntactic theory. The two types of approaches are in the same boat; the question for all of them is what lexical entries look like, and what information is stored where.

[15] Personally, I suspect that the amount of idiosyncrasy in word formation has been vastly overstated. Cases like profess—professor are few and far between, and most English speakers probably consider the two words to be unrelated. The word profession can also be used to mean ‘the act of professing’, as in professions of faith, which is completely regular. To my knowledge, no one has actually done a systematic analysis to back up the claim of massive idiosyncrasy, or to compare the amount of lexical idiosyncrasy to the amount of idiosyncrasy of phrases. Jackendoff (1997, 156) and Sag et al. (2002, 2) estimate that the number of “multi-word expressions” in a speaker’s lexicon is at least as large as the number of single-word expressions.

[16] Reinhart and Siloni (2005) also claim that it is no accident that the lexicon languages do not permit reflexivization of raising to object verbs: in their system, this follows from the operation applying in the lexicon, where only selected arguments of a verb stem can be affected. However, it also follows from these languages limiting reflexivization to a small, listed set of verbs; no raising to object verb is on that list. The list only consists of grooming verbs, none of which are raising to object verbs.
If we give up lexical rules and only have a single syntactic rule component, as I am arguing for here, then all word formation has to be done by the same syntax that builds phrases. We need a way to capture idiosyncrasies and lack of full productivity. There are two obvious ways:

1. All specifications are stored in the lexical entries of roots. Suppose we have a root SOCI. In the entry for this root, it can be recorded that SOCI + -al = ‘pertaining to society’, while SOC + -al + -ist = ‘follower of the doctrine of socialism’, and so on. (Cf. Marantz 1997)

2. Lexical entries include syntactic structure. There is a lexical entry for socialist that includes its structure and its meaning. There is another lexical entry for social that includes its structure and meaning (and its structure is probably a subset of the structure of socialist). (Cf. Hale and Keyser 1993)

On both views, combinations are memorized as they are encountered. At the same time, language users will generalize to some extent and extract commonalities (different individuals may do this to different extents). This is not really different from most lexicalist views; all that is different is the claim that the system that puts words together is the same as the system that puts phrases together.

There are also other possibilities besides the two listed, and I will not commit to any one here. It is an empirical question what the best account is. It seems to me, however, that a uniform account of idiosyncrasy at the word level and at the phrasal level would be desirable, and this is more likely to be successful in a theory that does not distinguish the two. (See Bruening 2010 on an approach to phrasal idioms where selection is crucial; selection operates uniformly throughout the system. See also Sag 2007)

4.5 Summary of Sections 2–4

In sections 2–4 we have seen that the view of grammar proposed by the Lexicalist Hypothesis is incompatible with the facts of numerous languages. There is no strict separation of the component of grammar that produces words and the component that produces phrases. Phrasal syntax can feed word formation processes, and there are processes that can target both phrasal units and sub-word units. Additionally, principles that have been claimed to distinguish word formation and phrasal syntax actually do not.

So far, then, there is no argument for strictly separating the word and phrase systems, and doing so leads to incorrect predictions.

5 The Lexicalist Hypothesis is Superfluous

I now show that we do not need the Lexicalist Hypothesis to account for the facts that it is meant to account for. We saw in the last section that the Lexicalist Hypothesis is wrong in numerous ways, but there are other cases where it does seem to capture something. That is, there are cases where phrasal elements cannot target sub-word units, and there are cases where sub-word units seem to behave differently from their phrasal correspondents. I show here that these facts have independent explanations, with no need to invoke the Lexicalist Hypothesis or the Principle of Lexical Integrity (Bresnan and Mchombo 1995). An important point in this section will be that the constraints that we see actually make no reference to the notion of a word. We saw this already concerning phrases inside words in section 2.1: the relevant phenomenon occurred inside words as well as across word boundaries. We will see this again here: the units that are relevant are not words. This means that the Lexicalist Hypothesis, which crucially picks out words, does not actually account for anything.

I begin in the first three subsections with the arguments that Bresnan and Mchombo (1995) give for their Principle of Lexical Integrity. This principle says that word parts are inaccessible to the syntax, as would follow...
from the Lexicalist Hypothesis. We saw in section 3 that this is incorrect in many instances, but in others it does seem to be true.

5.1 “Lexical Integrity”: Extraction

The first argument involves extraction. According to Bresnan and Mchombo (1995) word parts are inaccessible to extraction. I illustrate here with parts of compounds:

(61) * It’s American history that they’ve been [— teachers] for years. (modified from Bresnan and Mchombo 1995, 187, (3b))

Presumably, if there were no Principle of Lexical Integrity, we would expect word parts to be accessible to extraction.

The obvious question to ask here is whether the phrasal syntax independently rules out such attempts at extraction. It does, in fact. Extraction, and A-bar extraction in particular, may only target phrases, and not heads (cf. Haspelmath 2011, 52):

(62) a. It’s [a dedicated teacher] that she’s been — for years.
   b. * It’s teacher that she’s been [a dedicated —] for years.

(63) a. As [incredibly terrified of crocodiles] as you’ve been — over the years, . . .
   b. * As terrified as you’ve been [incredibly — of crocodiles] over the years, . . .

Neither of these illicit instances of extraction is ruled out by the Principle of Lexical Integrity. Independently, then, we need a constraint to the effect that A-bar extraction only targets phrases, and may not target heads. Once we have this constraint, however, the inability of A-bar extraction to target sub-word units follows: they are not phrases. There is no need to have a Principle of Lexical Integrity in addition. It is entirely superfluous.

The same constraint operates on A-extraction, too. Raising to subject and passivization may not target a head, even in a case like the prepositional passive where a subpart of the argument of the verb—the complement of the P complement of the V—can be extracted:

(64) a. The wealthiest candidate is likely — to win.
   b. * Candidate is likely [the wealthiest —] to win.

(65) a. Only the softest bed will be slept [in —].
   b. * Bed will be slept [in [only the softest —]].

Both A- and A-bar extraction are constrained to operate on phrases, not heads; and it therefore follows that they cannot target sub-parts of words, because those are not phrases.

Some lexicalist theories treat raising and passive lexically (e.g., Bresnan 1982, Müller 2006). As such, they should in principle be able to target sub-parts of words, yet they cannot:

(66) a. * American history is likely [a(n) — teacher] to win Teacher of the Year.
   b. * American history was hired [a(n) — teacher].

Some other principle must block this. Typically, this is done simply by the way the rule is stated: it targets an argument. A sub-part of an argument cannot be targeted. But this by itself rules out targeting a sub-part of a word, since a sub-part of a word is not an argument. The Principle of Lexical Integrity is thereby rendered superfluous even within a lexicalist theory. It does no work whatsoever regarding extraction, even in a theory that assumes its correctness.

There are proposals that A-bar extraction can target heads, for instance Müller’s (1996) analysis of partial VP fronting in German. However, it is not necessary in these analyses that A-bar extraction be able to target a head. Partial VPs can also be targeted, and a verb by itself can be a partial VP. That is, phrases can contain only one word and still be phrases.
Since I am going to propose below that the head-phrase distinction plays a large role in explaining other facts that form the “residue” of the Lexicalist Hypothesis, it is important to elaborate on this distinction here. First, it is important to make it clear that the syntactic notion of a head is not the same as a word. In X-Bar Theory, a head is a zero-level category. Words can consist of multiple heads. Heads can also be complex, and themselves consist of multiple heads. Conversely, it is possible that the pieces of a phonological word may not have to all be dominated by a single X-zero node in the syntax (for instance, shouldn’t’ve might span multiple head positions; note that it cannot front as a unit in subject-auxiliary inversion, which is commonly viewed as head movement). I will not attempt to define the notion of a word (something that might be impossible; see note 2). As for the notion of a head, I will simply assume that there is such a notion, and it is what is typically meant in syntactic theories that adopt some form of X-Bar Theory. As a rough guide, most morphemes are heads, and a complex of morphemes is a complex head (which may or may not be a word). It is not my intent to spell out a complete syntactic theory of word formation here. An intuitive understanding of what a head is—something that is common to most syntactic theories—is all we need to explain the phenomena discussed in this section. What is important is that a head is not the same as a word, and so I have not simply reintroduced the Lexicalist Hypothesis under a different guise. For instance, the ban on extracting heads rules out extracting both words and sub-parts of words. This is very different from the Principle of Lexical Integrity, which only bans extracting sub-parts of words.

To illustrate further with extraction, consider compounds again. We saw in section 2.1 that the first member of a compound can be a syntactic phrase. Such phrases still cannot be extracted by A-bar movement processes, however:

(67) * It’s itself that it has a [curled-up-on — look].
   (cf. It’s itself that it has curled up on.)

Any theory of compounds has to say that the phrase curled up on itself is turned into a noun stem, as described in section 2.1. As also discussed in that section, the internal structure of stems in general is invisible to elements outside that stem. As soon as the stem \[N \[XP curled up on itself\]] is created, then, the internal structure of N becomes invisible. This is due to strict cyclicity, the correct form of which should cover both morphology and syntax since they are the same (but I will not attempt to spell out a theory of cyclicity here). So, the fact that there is a phrase inside the compound is not something that can be accessed in the syntax when it goes to do extraction. The XP inside the N head is simply invisible, as are all the phrases that it dominates. (For an analysis of phrasal compounds that attempts to derive their inaccessibility to extraction from cyclicity, see [Sato 2010].)

Note that it is not important to this explanation that the phrase is inside a word. It is important that it is inside a head. It is also important to note once again that strict cyclicity in morphology does not follow from the Lexicalist Hypothesis, since it holds within as well as across words. Thus, this explanation is not simply rephrasing the Principle of Lexical Integrity, and the Lexicalist Hypothesis really can be done away with, in favor of independently needed constraints that do not refer to words.

5.2 “Lexical Integrity”: Conjunction and Ellipsis

According to [Bresnan and Mchombo (1995)] conjunction and ellipsis are phrasal processes and so cannot target parts of words. We saw in section 3.1 that this is false. Word parts can be conjoined and elided. Above we saw that all instances of conjunction of word parts might actually be ellipsis, however, so it may well be that conjunction cannot target word parts. If this is correct, do we need a Principle of Lexical Integrity to explain it?

The answer is no. Again, all we need is the recognition that some syntactic processes may only target phrases and not heads. Conjunction may be such a process. Of course, conjunction does seem to be able to operate on heads, as in the following examples:

(68) a. You can bring this or that water bottle.
    b. They can and will arrest you.
    c. They shot and killed the suspect.
There are some indications that coordination of heads is illusory, however, and is probably always phrasal coordination (as proposed by Kayne 1994 and Beavers and Sag 2004). For instance, coordination of demonstratives and modals can include phrasal elements like adverbs and negation, and in coordination of verbs, and can always be replaced by but did not:

(69) a. You can bring this but probably not that water bottle.
    b. They can and probably will arrest you.
    c. They shot but did not kill the suspect.

That is, any apparent instance of head coordination can include more than one word in one of the conjuncts, making it phrasal.

Similarly, coordination of verbal heads (here nominalized) acts like coordination of word parts in permitting plural agreement and antecedence of reciprocals (see above and Chaves 2008):

(70) a. Their shooting and killing (of) the suspect were unrelated to each other. (the suspect was wounded by the shooting last year, and then they killed him with a knife this year)
    b. Her offering and making me an espresso usually take place on different days.

It is plausible, then, that all instances of apparent head coordination are actually phrasal coordination with ellipsis, or are instances of right node raising (as Chaves 2014 seems to analyze them).

Borsley (2005) and Abeillé (2006) argue against this view, arguing that there must be true coordination of heads. However, their arguments do not go through. In every case where they contend that some phenomenon distinguishes head coordination from phrasal coordination (primarily right node raising), the same fact holds with what must be phrasal coordination (but did not). As an example, Abeillé (2006) claims that head coordination and right node raising differ in available interpretations. According to her, head coordination in (71a) and right node raising in (71b) differ in that in (71a), there are necessarily only two books, whereas in (71b) there could be either two or four books:

(71) a. Paul read and annotated two linguistics books.
    b. Paul read, and Mary annotated, two linguistics books.

This is true, but a variation on (71a) that is clearly phrasal (72) also obligatorily involves only two books:

(72) Paul read but did not annotate two linguistics books.

Example (72) simply cannot be analyzed as head coordination, since the second coordinate involves three different words. The forced “single object” interpretation could therefore not be explained by the coordination being head coordination.

The same is true of all of the differences Borsley (2005) and Abeillé (2006) point to between right node raising and apparent head coordination: all the respective facts still hold when and is replaced with but did not. For instance, weak pronouns are still allowed in (73b), just as they are in (73a), although they are not very good in canonical right node raising (73c):

(73) a. They shot and killed him.
    b. They shot but did not kill him.
    c. ?? The police shot, and the gang members finally killed, him. (odd without heavy stress on the pronoun)

For further discussion of prosody in right node raising, see Chaves (2014). Prosody does not argue for the existence of head coordination, because the same prosody of (73a) holds in examples like (73b), which is clearly phrasal.

Additionally, in all the cases from other languages cited by Abeillé (2006), the phenomena claimed to be limited to heads also always admit “light” modification of the putative head, making it necessarily phrasal. Abeillé’s own analysis appeals to phonological weight, and it is likely that that is the factor involved in her data, not the head-phrase distinction.
I conclude that the arguments for the existence of head coordination do not go through, and there are good reasons to think that, in fact, apparent coordination of heads is always coordination of phrases.\footnote{19} If this is true, then coordination is just one of many processes (like extraction above) that only targets phrases, and not heads. This is a restriction in the phrasal syntax, but it has the result that the coordination of word parts is banned, with no reference to the Lexicalist Hypothesis or the Principle of Lexical Integrity. Once again, all the work of the Lexicalist Hypothesis is already done by the phrasal syntax.

Although coordination is limited to phrases, ellipsis clearly is not (see section 3.1). Neither is focus (section 3.2). Ellipsis and focus seem to be able to target any syntactic unit: morphemes, complex heads, phrases. (Focus also seems to be able to cross cyclic boundaries, but in the phrasal syntax it is known to be insensitive to islands.) In the typology of syntactic elements, then, we have ones that can only target phrases (extraction, coordination); ones that can only target heads (some morphemes discussed below); and ones that can target anything (ellipsis, focus). Note that this typology makes no reference to the notion of a word.

5.3 “Lexical Integrity”: Inbound Anaphoric Islands

Bresnan and Mchombo (1995) also cite “inbound anaphoric islands” as evidence for their Principle of Lexical Integrity. Following Postal (1969, 213-214) they note that words may be formed from referential nouns, but not from pronouns.\footnote{20}

(74) McCarthyite, *himite

Other types of pro-forms also cannot appear inside words:

(75) a. * People who smoke like other do-so-ers. (Postal 1969, 217, (69a))

b. * We need a truck but not a one-driver. (cf. We need a truck but not a driver of one.)

It is not exactly clear how this would follow from the Principle of Lexical Integrity. According to Simpson (1983a) coreference is part of the sentence grammar (the phrasal system), so a pronoun could not receive its reference within a word, since the sentence grammar has no access to sub-words parts. As Ward, Sproat, and McKoon (1991) note, this would then have nothing to say about cross-sentential anaphora, which is also ruled out. Bresnan and Mchombo (1995) themselves explain the facts in a way that makes no reference to Lexical Integrity: “indexical pronouns, though they do have intrinsic lexical content and can appear word-internally, lack the appropriate lexical content to serve as morphological bases for semantic derivatives” (p192). In other words, pronouns have very little semantic content, and that is what stops them from forming words like *himite and *do-so-er. In attested words with pronouns like he-man and she-wolf, the pronouns are used just for their gender features, which is the only real semantic content that they do have. If this is correct, then inbound anaphoric islands have nothing to do with the Principle of Lexical Integrity and everything to do with the paucity of content in pro-forms.

A different explanation is offered by Sproat (1988). According to Sproat, pro-forms are always maximal projections. Pronouns are actually NPs, do so is a full VP, and so on. Word formation processes, according to Sproat, never operate on maximal projections. For instance, it is possible that suffixes like -ite and agentive -er strictly select heads, and may not combine with phrases (see more on this below). This rules out formations like *himite and *do-so-er.

Note that either of these potential explanations would permit pro-forms in compounds. In Sproat’s theory, pro-forms should be allowed if a phrase can form the input to a word, and we know this is possible in a compound.\footnote{25} Postal (1969) also discusses “outbound anaphoric islands,” but these are not considered by Bresnan and Mchombo (1995) to be evidence for Lexical Integrity, since they were shown by Ward, Sproat, and McKoon (1991) to be regulated by pragmatics.
In Bresnan and Mchombo’s analysis, if a pro-form can be contentful, it should be allowed, and pro-forms can get content in phrases like those that form the input to compounding. It is correct that phrasal compounds can include pro-forms, including fully referential ones:

(76) a. The dog gave me an accusing you-like-her-better-than-me look.
    b. She gave the flowers a resigned why-do-they-always-buy-me-ugly-ones look.
    c. When he said, “go ahead and dance,” she gave him a scornful I-won’t-do-so-alone look and he dutifully stood and accompanied her to the dance floor.

Recall that Bresnan and Mchombo (1995) claim that phrases in compounds are “lexicalized” (see above). If they are lexicalized then they would be expected to behave like words, and be inbound anaphoric islands. Instead they behave like phrases, and permit pro-forms.

Additionally, Bresnan and Mchombo (1995) themselves cite numerous counterexamples to the inbound anaphoric island constraint, where in some languages a derived word form can include pronominal agreement affixes. Harris (2006) similarly shows that Georgian word-formation processes can include fully referential pronouns. This means that inbound anaphoric islands are not observed in some languages (or in compounds in English), a state of affairs that should be impossible if they really follow from the putatively universal Principle of Lexical Integrity.

Inbound anaphoric islands, then, follow the pattern we have seen throughout this paper: there are numerous counterexamples to what the Lexicalist Hypothesis requires, and where there are restrictions, they follow from other considerations.

I turn now from the Principle of Lexical Integrity to other claims that have been made in the literature.

5.4 Chomsky’s “Internal Structure” Argument

One of Chomsky’s (1970) original arguments for a split between a word system and a phrase system was the claim that derived words have the phrasal structure of their derived category, and do not have the phrasal structure of the category they are derived from. For instance, a deverbal nominalization has the same phrasal syntax as an underived noun, and does not have the phrasal structure of a verb or a sentence. Newmeyer (2009) illustrates this with the following examples:

(77) (Newmeyer 2009, 95, (9))
    a. the stupid refusal of the offer
    b. * the refusal stupidly of the offer
    c. * the not refusal of the offer
    d. * the have refusal of the offer

(78) a. * She stupid refused the offer.
    b. She refused the offer stupidly.
    c. She did not refuse the offer.
    d. She has refused the offer.

A derived N takes adjectives, not adverbs, and may not have sentential negation or auxiliary verbs that are possible with the verb it is derived from. According to Chomsky and Newmeyer, these facts only follow if derived Ns enter the syntax as unanalyzable Ns.

Chomsky (1970) originally contrasted nominalizations with gerunds, which have the external distribution of nouns but the internal structure of VPs or even clauses:

(79) a. * her stupid refusing the offer
    b. her stupidly refusing the offer
    c. her not refusing the offer
    d. her having refused the offer
According to Chomsky (1970), gerunds are derived transformationally; if nominalizations are not, we explain the differences between them.

However, later work in lexical frameworks like LFG and HPSG has lost the ability to make such a distinction. In these approaches, any process of word formation has to be a lexical process (Dowty 1978, 412, Bresnan 1982, 21, Müller and Wechsler 2014, 32). Since gerunds are morphologically derived forms, they must be lexical, too. An HPSG analysis of gerunds is proposed in Malouf (2000); an LFG analysis is proposed in Bresnan and Mugane (2006). Their properties are captured with lexical rules and constraints in these analyses.

This means that the Lexicalist Hypothesis again does no work in explaining linguistic facts: all the work is done by the actual analysis. Nominalizations are analyzed one way, gerunds another; their differences are captured by analyzing them differently. The Lexicalist Hypothesis plays no role whatsoever.

Moreover, the claim that Ns derived from Vs do not have phrasal properties of Vs has been argued to be false, although the facts are disputed. Fu, Roeper, and Borer (2001) argue that Ns derived from Vs may have adverbs following them, while underived ones may not:

(80) (Fu, Roeper, and Borer 2001, 549, (2))
   a. The occurrence of the accident suddenly disqualified her.
   b. * Kim’s accident suddenly on the track disqualified her.

This has been contested by Newmeyer (2009), but all of his examples of putatively underived nouns (p.109, examples 47a–c) are either zero-related to verbs (use, release) or can plausibly be analyzed as derived from a verbal source that is not actually used (recourse, which is used in the “light verb” construction take recourse). The following contrast between a derived N and an underived one appears to me to support Fu, Roeper, and Borer’s view over that of Newmeyer:

(81) a. The sudden growth of the tomatoes really shocked me.
   b. The growth of the tomatoes so suddenly really shocked me.
(82) a. the overwhelming smell of rotten fish
   b. * the smell of rotten fish so overwhelmingly
   c. * the smell so overwhelmingly of rotten fish

However, I acknowledge that the matter requires more research to be settled. (Newmeyer’s skepticism of the do so data also offered by Fu, Roeper, and Borer 2001 does appear to be warranted.)

Other facts have been suggested to distinguish derived from underived Ns. For example, Bruening (2013) argues that certain PP adjuncts require verbal structure, and so they are only permitted with Ns that include verbal structure. Instrumentals are one such. They are allowed with VPs and with derived nouns, but not with (at least some) underived nouns (Bruening 2013, 12, (48–52)):

(83) a. The inspector saw the blood with a microscope.
   b. The sample was smelled with an electronic nose.
   c. The danger was sensed by Peter with his spider-sense.
(84) a. the perception of light with a photosensor
   b. the detection of the sound with an amplifier
   c. the discernment of God’s will with various omens
(85) a. * the sight of the blood with a microscope
   b. * the smell of the sample with an electronic nose
   c. * Peter’s sense of danger with his spider-sense

Ackema and Neeleman (2002, 119, (41)) present some examples of adverbs with apparently underived nouns as grammatical, but in my judgment they are terribly ungrammatical. The same judgments are repeated in Ackema and Neeleman (2004, 22, (9))

27
We also saw in section 2.3 that nominalizations can include resultative secondary predicates in German. If resultative secondary predicates can only combine with VPs (something that seems to be true in English), then these nominalizations include verbal structure.

At the same time, it is true that nominalizations differ from gerunds in having largely nominal syntax. Any theory will have to account for this difference, whether it assumes the Lexicalist Hypothesis or not. As described above, lexicalist theories do this by analyzing gerunds and nominalizations as both lexical, but treat them differently. A purely syntactic theory would derive them both syntactically, but again would have to treat them differently. Perhaps the simplest analysis would treat them differently only in size: nominalizations nominalize a fairly small structure, say VoiceP as in [Bruening 2013, 31–34, while gerunds nominalize a larger phrase, perhaps a full IP as in Abney 1987]. Producing an analysis is not important here; the important point is that lexicalist and syntactic theories are entirely equivalent on this point. All the work is done by the actual analysis; no work is done by the Lexicalist Hypothesis.

5.5 Phrasal Elements That Cannot Target Sub-Word Units

According to [Williams (2007)] we need the Lexicalist Hypothesis in order to account for the inability of wh-questioning to target part of a word, as in the following:

(86) [Williams 2007, 354, (1)]
   a. How complete are your results?
   b. * How completeness do you admire? (intended: [how complete]-ness do you admire)
   c. What degree of completeness do you admire?
   d. How complete a record do you admire?

In (86b), how cannot modify just the complete part of the noun completeness. The examples in (86c–d) are meant to show that there is nothing semantically ill-formed about this question.

I will address this issue in two parts. First, we need to know why how cannot target just a sub-part of a word. That is the topic of this subsection. Then, we need to rule out the bracketing in (86b), where the suffix -ness attaches to a phrase. That is the topic of the next subsection.

We saw in section 3 that certain processes can target sub-word parts. In particular, ellipsis and focus can. Apparently, a wh-phrase like how cannot. The question is whether we need the Lexicalist Hypothesis to explain why it cannot.

The answer is no. Since at least Bresnan (1973) and Jackendoff (1977) degree elements like how and quite have been treated as elements that are sister to a phrase (for instance, they are specifiers in X-bar theory: for a recent HPSG treatment of how as a specifier, see Ginzburg and Sag 2000, 187–188). This by itself rules out having them modify a sub-word unit: the only thing they can modify is the constituent X-bar (or some other phrasal constituent):

22One argument presented against syntactic accounts of nominalizations is that they can be coordinated with underived nouns and share arguments ([Wechsler 2008a, Müller and Wechsler 2014]). In section 5.2 I suggest that all apparent coordination of heads is actually coordination of phrases. If this is correct, the possibility of coordination of derived and underived nouns is not problematic for any theory. Additionally, the argument relies on the assumption that only heads with the same number and type of arguments can be coordinated, but this is false:

(i) a. She described and (then) made me a wonderful espresso.
   b. * She described me a wonderful espresso.

(ii) a. She recommended and (then) brought me a slice of key lime pie.
   b. * She recommended me a slice of key lime pie.

This is possible even with weak pronouns (she described and then brought me it), which is supposed to rule out a right node raising analysis (see more on this in section 5.2). The fact that derived words can be coordinated with underived words shows nothing.
How cannot be the specifier of an NP (*how man, *how liberty), so it simply cannot modify the N completeness. As a specifier, it also cannot attach directly to an A head, which might then permit further suffixation. We need some such constraint independently, to rule out cases like the following:

(88) a. How insufferably stupid is he being today?
   b. * Insufferably how stupid is he being today? (*[AP insufferably [A how [A stupid]]])

(89) a. How justifiably angry is he?
   b. * Justifiably how angry is he? (*[AP justifiably [A how [A angry]]])

As can be seen, how only combines with a full phrase, including any adverbs, and may not combine just with a head.

The analysis of how as a specifier captures this (assuming adverbs can only attach below thespecifier). But this analysis again rules out how combining with a sub-part of a word, with no need for the Lexicalist Hypothesis.

5.6 The “No Phrase Constraint”

The question now is what rules out the bracketing in (86b), repeated below, where how combines with an A before the suffix -ness turns it into a noun.

(90) * [N [AP how complete] -ness]

Similarly, Bresnan and Mchombo (1995) say that the Lexicalist Hypothesis and their Principle of Lexical Integrity are needed to account for the inability of adverbs and other phrases to modify sub-word parts:

(91) (Bresnan and Mchombo 1995, 192, (17))
   a. [A happy]-ness
   b. * [AP quite happy]-ness
   c. * [AP more happy [than sad]]-ness

All of these are supposed to be ruled out by something like the No Phrase Constraint (Botha 1981) in the Lexicalist Hypothesis: phrases may not form the input to word-formation. (In most versions this just follows from the hypothesized architecture of the grammar and need not be stated as a separate constraint; see note 3.)

At various points above, I have suggested that certain affixes might actually be able to combine with phrases (importantly, phrases that have not been turned into heads as in compounding). One example was the passivizing affix, which can apply to a V-PP phrase. Depending on how the data turn out, nominalizing affixes might be another example. If some affixes can combine with phrases, then what rules it out in this particular case?

23 A reviewer points out that the German equivalent of how seems to have different properties. I will leave exploration of other languages to future research, but again it is important to emphasize that we need detailed analyses of every individual morpheme, as they do not act alike.

24 It should be noted that it is possible for how to question a sub-part of a word, but that sub-part needs to be repeated and seems to be treated as a full phrase:

   b. How ‘post’ does ‘postcolonial’ have to be before it ceases to be a—or the—primary determinant in the way Irish writing is read, and reads itself? (Irish Poetry Since 1950: From Stillness Into History by John Goodby, p319)

I will leave an analysis of these types of examples to future work.
One type of analysis that has been proposed within a purely syntactic approach to word formation is to hypothesize that morphemes select different categories to combine with. For instance, Pylkkänen (2008) proposes that causative morphemes in the world's languages divide into three types. One selects a full phase (in the phase theory of Chomsky 2000), one selects a verb phrase, and one selects a bare root. The first two combine with phrases and may include things like adverbs. The root-selecting causative morpheme combines only with a bare root, and not a phrase.

Stating this slightly differently, and in a way consonant with the typology of linguistic elements proposed in section 5.2, one could propose that some morphemes select only heads, and may not combine with phrases. (Heads, unlike roots, can be complex.) The morpheme -ness would be one such morpheme. (The analysis of -ness in Embick and Marantz 2008 seems to have this character.) In support of this sort of analysis, complements of adjectives are generally not preserved from the adjective to the noun in -ness:

(92) a. We are mindful of each other.
   b. * our mindfulness of each other
(93) a. We are worthy of each other.
   b. * our worthiness of each other
(94) a. We are crazy about each other.
   b. * our craziness about each other
(95) a. We are fond of each other.
   b. * our fondness of each other

A complement can be preserved just if it is incorporated, forming a complex head:

(96) credit-worthiness, (certificate of) road worthiness, boy-craziness, ...

Note that the morphemes suggested to attach to phrases, the passivizing and nominalizing suffixes, do often preserve complements:

(97) Adjectival Passives
   a. We stuffed the pillow with feathers, The pillow remained stuffed with feathers. (based on Levin and Rappaport 1986, 634)
   b. She convinced me of her trustworthiness, I remain unconvinced of her trustworthiness.
   c. They based the character on George W. Bush, The character appears based on George W. Bush.
(98) Nominalizations
   a. destroy NP, destruction of NP (fully general)
   b. depend on, dependence on; rely on, reliance on
   c. apply for, application for; prepare for, preparation for
   d. believe in, belief in; specialize in, specialization in; succeed in, success in
   e. commit to, commitment to; refer to, reference to; respond to, response to
   f. approve of, approval of; acquit of, acquittal of
   g. agree with, agreement with; collide with, collision with

One could argue that there is a correlation between preserving arguments of the stem and being able to combine with phrases. If a stem must project its (internal) arguments in its minimal phrase, then any affix that combined with a phrase rather than just a head would preserve (internal) arguments of the stem it attaches to. An affix that attaches only to a head may or may not preserve the arguments of that head. Since the suffix -ness does not preserve arguments, we can conclude that it does not attach to phrases.
If this is true, then there is an independent explanation for the inability of \textit{-ness} to attach to APs with wh-phrases and adverbs, or phrasal comparatives. We do not need the Lexicalist Hypothesis to explain this, it follows from properties of individual affixes.

To clarify, what I am proposing is that individual affixes can differ in their selectional requirements. Some, like the passivizing affix, can combine with a whole VP:

(99) \[ \text{PassP} \]
\[ \text{Pass} \quad \text{VP} \]
\[ \text{V} \quad \text{PP} \]

Others, like \textit{-ness}, can only combine with a head, although that head can be complex and need not be a word:

(100) \[ \text{N} \]
\[ \text{A} \quad \text{-ness} \]
\[ \text{mind} \quad \text{-ful} \]

In the case of the passive morpheme, some mechanism is necessary to combine the morpheme with the head of its complement. This could be movement (e.g., \textit{[Baker 1988b]}) or it could be some other mechanism. The two could combine phonologically and not syntactically, for instance. Or, the form of the V could be determined to be its passive participle form because it occurs in the local context of Pass (so the actual phonological content of Pass never occupies the Pass head). There are numerous possibilities, which I will not decide between here. What is important is that the two types of affixes differ, and the effects of the No Phrase Constraint (where they do occur) fall out from the difference.

A more general point to be made here is the same one that was made earlier regarding \textit{self-} versus \textit{re-} (section 4.2): It is necessary to analyze individual affixes or word-formation processes in detail. They differ significantly from each other, in such a way that blanket statements like “affixes cannot attach to phrases” and “affixes may not take scope over adverbs” are simply false: some can, and those that cannot are often even more constrained, in peculiar ways. We do not want a model of grammar whose architecture rules out affixes combining with phrases and affixes taking scope over adverbs, because many of them do. If some of them do not, then that is something that is going to have to be captured in the analysis of those particular affixes, since it will not follow from general principles. I have given one suggestion for an analysis here, but it may well be that a better analysis is available. One can only decide this based on extensive research into the properties of the individual affixes involved.

5.7 Summary

This section has gone through several cases from the literature where some phenomenon has been claimed to require the Lexicalist Hypothesis and its Principle of Lexical Integrity. In every case, the Lexicalist Hypothesis is superfluous: all of the data are already accounted for by independently needed principles, even within theories that assume the Lexicalist Hypothesis. Importantly, these principles never refer to the notion of a word. Rather, they refer to syntactic units like ‘head’ and ‘phrase.’ Another important point is that we need detailed analyses of particular morphemes, since they can be both more and less constrained than what the principles of the Lexicalist Hypothesis require. The Lexicalist Hypothesis has no explanatory value in understanding any of the phenomena discussed.

6 Conclusion

The Lexicalist Hypothesis was a reasonable hypothesis about the organization of the human language faculty. It could have been correct that the human grammar has distinct word formation and phrasal syntax components. However, all of the evidence reviewed here indicates that it does not. The first part of this paper showed that there
are numerous phenomena where phrasal syntax provides the input to word formation, and there are also phenomena where phrasal operations have access to sub-word units. Claims that processes of word formation and processes of phrasal syntax obey different principles were shown to be incorrect. The second part of the paper showed that where there are facts to be explained, they follow from independent principles that make no mention of the word as a linguistic unit. Since the Lexicalist Hypothesis refers specifically to the notion of a word, it is incapable of explaining the facts that it was designed for.

If some hypothesis is both incorrect and does no work, the obvious action to take is to discard it. In this case, doing so leads to the kind of theory that considerations of parsimony would also prefer: a theory where there is only one combinatorial system for both words and phrases, not two distinct systems. As noted in the introduction, a theory with only one component is simpler than and therefore preferable to a theory with two components, all other things being equal. This paper has attempted to show that all other things are equal: the Lexicalist Hypothesis does no work, and so has no empirical advantage to outweigh the parsimony consideration. In fact, it seems to be at a disadvantage, since it gets many empirical facts wrong.

References


Shibatani, Masayoshi (1990), The Languages of Japan. Cambridge: Cambridge University Press.


