

Extending the Extension Condition to Discontinuous Idioms

Peter Svenonius
*CASTL, University of Tromsø**

September 27, 2005

Abstract

This paper makes a specific proposal regarding the structure of idioms as they are listed in the lexicon, which has important ramifications. Specifically, I suggest that first Merge as well as internal Merge can target subconstituents of structures already built by Merge. This creates structures which I call Banyan trees, in which there is more than one undominated node. I suggest that this model, together with certain assumptions about phases and about the functional sequence of syntactic heads derive important facts about the structure of idioms.

Keywords: idioms, Merge, Extension Condition, Banyan trees, phrase structure, constituency

1 Introduction

This paper is about idioms, by which I mean listed syntactic structures which interact with syntax in the way words do, and have unpredictable meanings in the way that words do, but which consist of more than one “piece.” This is roughly the usual sense of idiom, if by “piece” we mean ‘phonological word.’ If, on the other hand, “piece” is taken to mean ‘syntactic head,’ then this is the sense of idiom used by Marantz (1997a). For the purposes of this paper, I will draw on evidence from multi-word idioms

*Thanks to Gillian Ramchand, Jacek Witkoś, Noam Chomsky, and an anonymous reviewer named David Adger for feedback. Thanks also to seminar participants in Tromsø in spring 2005, and to audiences in Umeå and Poznań in April 2005, and at MIT in July 2005.

like *spill the beans*, though my conclusions are compatible with those of Marantz.

An idiom can generally be distinguished from a metaphor by the fact that there are specific words in an idiom which are necessary for the idiomatic sense; for example, the idiomatic sense of *spill the beans*, namely ‘divulge the secret,’ is lost if either of the words *spill* or *beans* are absent, e.g. in *spill the peas* or *pour out the beans*. In contrast, the metaphoric sense of *bang your head against a problem*, namely ‘try hard, in vain, to solve a difficult problem,’ remains when *bang* is replaced with near-synonyms such as *bash*, *dash*, *smash*, *pound*, *hit*, or when *head* is replaced with *noggin*, *skull*, *brain*, and so on.

An idiom can be used as the basis for a metaphor, so that the idiom *beard the lion in his den*, meaning ‘confront an adversary on his own territory,’ might be used to construct sentences like those in (1).

- (1) a. I’ll beard the editor in his den.
- b. I’ll beard the lion in his office.
- c. I’ll interview the lion in his den.

Such examples allude to an idiom without actually containing one. I assume that the examples of idioms in control structures discussed by Ruwet (1991) and Nunberg et al. (1994)¹ are also metaphors which allude to idioms, though the distinction is subtle and more can certainly be said about this matter than I will here.

Jackendoff (1997) points out that in order to retain the idiomatic sense, a particular syntactic configuration is furthermore necessary; for example *the cat is out of the bag* means something like ‘the deception is over,’ or ‘the secret is out,’ but *the bag is out of the cat* has no such meaning, or even any related one. Jackendoff concludes that the syntactic structure must be listed as well as the individual vocabulary items participating in the idiom. He gives an example (his p. 162) like the one in (2) for the idiom *take (someone) to the cleaners*, ‘cheat (someone) of his or her money.’

¹For example, Nunberg et al. (1994) give examples like *Every lion prefers to be bearded in his den*.

- (2) a. _atake _bto _cthe _dcleaners
 b.

```

graph TD
  VP_x[VP_x] --- V_a[V_a]
  VP_x --- PP[PP]
  PP --- P_b[P_b]
  PP --- NP[NP]
  NP --- Det_c[Det_c]
  NP --- N_plur["[ N_d, plur ]"]
  
```


 c. [Event GET ([]_A, [ALL OF [MONEY OF []_A]])]_x

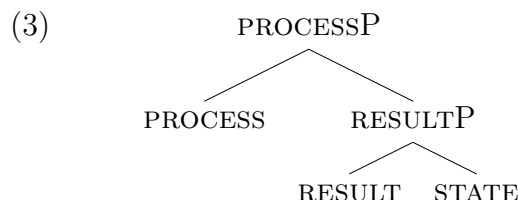
(2a) gives the vocabulary items which must be mapped onto the syntactic structure in (2b), according to the subscripts indicated, while (2c) gives the semantic interpretation (Jackendoff’s Lexical Conceptual Structure, or LCS). Linking rules (Jackendoff 1990) govern how the arguments in (2b) map onto the empty bracketted slots in the LCS in (2c), according to a thematic hierarchy defined over the primitives of the LCS. Jackendoff suggests that a TAG grammar (Abeillé 1995) splits the tree to rearrange it around the (open) object position.

(2) suggests that very specific structured information can be listed in the LCS of an idiom, and that the relationship of the syntax and the semantics is relatively loose. A more careful examination of the semantics of the idiom calls this into question, however; for example, it is not clear that the idiom involves universal quantification of the sort implied by the use of “ALL” in (2c); in fact, the meaning might be characterized as more vaguely resultative, for example “cheat thoroughly”; a single-word translation might be *fleece*.²

Marantz (1997a) argues that the specifics of Jackendoff’s solution go too far in separating meaning from form; the meanings of idioms do not violate general meaning-form correlations that hold over non-idiomatic, regularly composed structures. To take (2) as an example, it seems to fit rather well with the structure of compositional resultatives in English, like *flog* someone *to death* or *cheat* someone *to the poorhouse*. I assume, therefore, that the syntactic structure is that of a resultative, and therefore necessarily maps onto a resultative semantics, not by coindexing a tree with an idiosyncratically listed LCS, but by regular, deterministic rules that map that syntactic structure onto that semantic representation. Schematically,

²BBC World Service online (www.bbc.co.uk) suggests as a translation for *take someone to the cleaners* ‘deprive someone of their money or possessions,’ while *NTC’s American Idioms Dictionary* has the vaguer (Spears 2000) definition ‘to abuse or damage someone.’

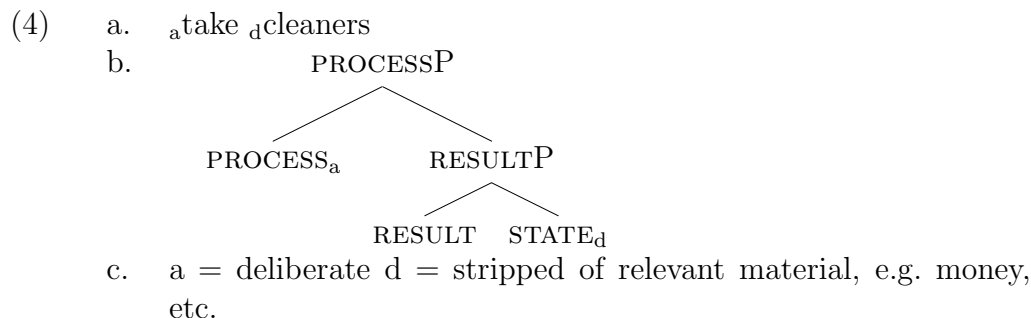
the syntactico-semantic representation might be something like that in (3) (see Ramchand and Svenonius 2002, Ramchand 2005 for further specifics regarding resultative structures).



The two parts PROCESS and RESULT (cf. Ramchand's 2005 *proc* and *res*) are primitive components of a wide range of verbs and can only combine in this way, much as D is standardly assumed to take N as a complement and not vice-versa.

Still, because of the nature of the idiom, at least the words *take* and *cleaners* have to be idiosyncratically linked to parts of the listed structure;³ possibly, the words *to* and *the* could be made to follow from syntactico-semantic features of the construction. Encyclopedic information, for example the fact that the end state is one of impecunity or thorough cheatedness, must also be listed.

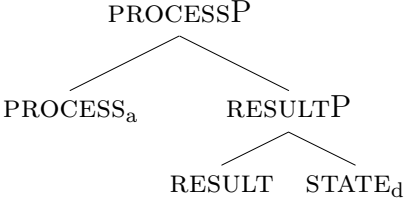
In this paper I will sketch an approach to idioms which should make it possible to greatly simplify the lexical entries for idioms, compared to Jackendoff's proposal, in the direction of something like (4).



Here, more of the syntactic structure is predictable than Jackendoff assumes, hence need not be listed; and there is no structured LCS distinct from the syntax, (4c) expressing only encyclopedic information (which might be much richer than I have indicated, but is crucially not structured in a syntactic way, apart from possible links from meanings to parts of the idiom, as indicated). Furthermore, I will not make use of TAG grammar to merge in

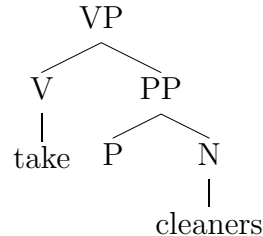
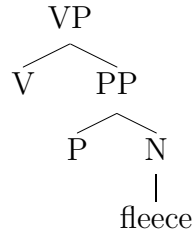
³As Jackendoff (1997) notes, it is the word *take* which is linked here, not some homophone of it, since the past tense is *took* and the participle *taken*.

the internal argument, though I will propose something similarly unconventional regarding the nature of Merge. The verb *fleece* might likewise have a lexical entry like that in (5), again borrowing Jackendoff's notational conventions.

- (5) a. _dfleece
 b. 
 c. a = deliberate d = stripped of relevant material, e.g. money, etc.

The verb has exactly the same syntax as the multi-word idiom. This is a quite general fact; idioms have meanings that single words can have, hence the same structures. Importantly, the structure that is listed here in (5b) as part of the lexical entry is built by regular syntactic rules; having been built, it can also be listed. No word or idiom, I assume, can have a structure that could not be built by regular syntactic rules.⁴

I will henceforth represent idioms in a kind of shorthand, using more or less traditional category labels for syntactico-semantic primitives and attaching phonological words to the nodes to which they are linked. (4) can accordingly be abbreviated as in (6a) and the entry for *fleece* in (5) as in (6b).

- (6) a.  b. 

The idea is that since no phonological material is associated with V in (6b), it cannot be filled by any lexical verb. The noun *fleece* therefore moves to this position to support tense morphology (merged outside the idiom), more or less as suggested for *dance* by Hale and Keyser (2002), though *dance*, being non-resultative, would have no P projection. In the sections to follow,

⁴Much-discussed examples like *by and large* ‘generally’ seem to repudiate this. However, consider the possibility that /baɪ/ in this example is an idiomatically listed adjective, only appearing in this idiom the way the noun *petard* only occurs in the idiom *hoist by his own petard* ‘damaged by his own attack.’

I will explain how determiners, modifiers, specifiers, and other material are merged into these simple structures. Importantly, this shorthand is possible here because of the deterministic mapping of the syntax onto a semantics; the assumption is that there can be no structured semantic representation distinct from (6) in the lexicon, only encyclopedic information.

2 Regularities of idioms

Regularities of idiom structure have long been noted, for example by Chomsky (1980; 1981), Marantz (1984), and others. In this section I run through some of the regularities that I suggest should be captured in any analysis, and which I attempt to capture in the one outlined here.

2.1 The Structure of Idioms. It has long been noted that idioms tend to correspond to syntactic constituents (Marantz 1984). Thus, for example, there are hundreds or even thousands of idioms in English corresponding to verb phrases, with free subject positions, a few of which are illustrated in (7).

- (7) a. *kick the bucket* ‘die’
- b. *pass muster* ‘meet with approval’
- c. *pass the buck* ‘shirk responsibility’
- d. *pass the hat* ‘request donations’
- e. *bring down the house* ‘stimulate an audience to an enthusiastic response’
- f. *call off the dogs* ‘stop threatening’
- g. *call the shots* ‘be in charge’
- h. *climb the walls* ‘become agitated with boredom’
- i. *lay an egg* ‘put on a lackluster performance’
- j. *gum up the works* ‘slow down or complicate operations’
- k. *miss the boat* ‘be too late’
- l. *fly the coop* ‘escape’

There are also a few idioms which include a subject, some of which are illustrated in (8).⁵

- (8) a. *Heads will roll* ‘people will be punished’
- b. *The shit hit the fan* ‘a seriously bad thing happened’

⁵I am setting aside proverbs like ‘Rome wasn’t built in a day,’ which do not interact with productive syntax, unlike the examples in (8). Nunberg et al. 1994:516, n. 29 furthermore suggest that collocations like ‘A good time was had by all’ to be distinct from idioms, as they have literal, compositional meanings, and I will follow them in setting those aside here, though not without trepidation.

- c. *All hell broke loose* ‘the situation became chaotic’
- d. *The cat is out of the bag* ‘the secret has become known’
- e. *The jig is up* ‘the deception has come to an end’

However, there are exceedingly few idioms which include the subject and the verb, to the exclusion of the direct object. An apparent example is *God damn* DP; but the sentence structure of that example is different from that of ordinary sentences, as demonstrated clearly by Quang (1971).⁶

Free object idioms which are more difficult to explain away are *something’s eating* DP (‘something’s bothering DP,’ cf. Bresnan 1982, where the progressive aspect seems to be important) and *A little bird told me* (used to avoid naming a source, e.g. of gossip), if *me* can be replaced by other DPs. However, their very rarity is quite striking, and so I concur with Marantz (1984) that a fundamental subject–object asymmetry should be recognized in the analysis of idioms.

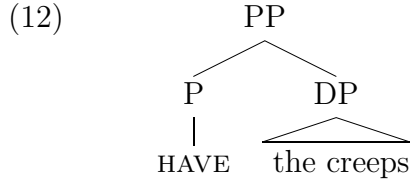
2.2 Free open positions. On the basis of such observations, it is often assumed that idioms are always syntactic constituents. Sometimes, there are plausible transformational derivations of discontinuous idioms like *give* someone *the creeps* ‘make someone uneasy,’ *give* someone *the sack* ‘fire someone,’ or *give* someone *flak* ‘chastise someone.’ Richards (2001a) points out the pattern in (9)–(11).

- (9) a. The morgue gives everyone the creeps.
b. Mary gave John the sack for showing up late.
c. Bill has given Susan flak about her behavior.
- (10) a. You get the creeps just from driving past the place.
b. John got the sack for showing up late.
c. Susan has gotten a lot of flak about her behavior.
- (11) a. *The creeps usually go away after you have worked there for a week.
b. *The sack has ruined many an employee’s Christmas.
c. *Flak only encourages her.

As Richards argues, the systematic alternation between (9) and (10) shows that neither *give* nor *get* is an obligatory part of the idiom. But the impossibility of the idiomatic reading in (11) shows that something that *give* and *get* share *is* part of the idiom. Richards’ solution is that the idiom contains an abstract verbal component HAVE, allowing these idioms to be

⁶E.g. as Quang points out, if *God* were the subject then **God damn himself* should be grammatical, and not *God damn God*, contrary to fact.

schematically represented as in (12).



If HAVE is a formative which must be supported by additional material in order to appear in a sentence, then we can assume that additional material must be merged on top of this structure, resulting in sentences like those in (9)–(10). Thus, these apparently discontinuous idioms are stored as constituents. The prediction is that idioms should not exist in which a verb like *give* (or its subcomponent HAVE) is part of the idiom along with the indirect object, while the direct object is an open position. This appears to be correct (also for Polish, cf. Witkoś and Dziemianko 2005).⁷

However, there are many idioms which do not correspond to traditional syntactic constituents, for example those with an open position in what is arguably an indirect object or small clause subject position.

- (13)
- a. *read DP the riot act* ‘give DP a stern warning lecture’
 - b. *drop DP a line* ‘contact DP’
 - c. *show DP the ropes* ‘give DP basic training’
 - d. *know DP from Adam* ‘know DP at all’ (NPI)⁸
 - e. *haul DP over the coals* ‘reprimand DP’
 - f. *serve DP right* ‘be a fair punishment for DP’
 - g. *send DP packing* ‘dismiss DP rudely’

Of course, there have been analyses, going back at least to Chomsky (1957), in which the object is interpolated into such sequences by transformation. But the evidence for such interpolation is generally less compelling in such cases than in the cases Richards discusses; e.g. the verb *read* seems to be part of the idiom in (13a), so the abstract HAVE analysis will not work straightforwardly there.

⁷Note the difference between (ia) and (ib), for speakers who make the distinction between stative *got* and inchoative *gotten*.

- (i)
- a. The cat has got/*gotten John’s tongue.
 - b. Susan has gotten/*got a lot of flak about her behavior.

This shows that the verb in (ia) is not the same verb, correctly predicting the unacceptability of **Mary gave the cat John’s tongue*.

⁸Negative Polarity Item (NPI): This idiom can only occur in the scope of negation or a similar operator.

Even more challenging for a transformational account is the type of idiom, very common in English, in which a possessor position is open, as illustrated in (14).

- (14)
- a. *bend DP's ear* 'talk a lot to DP'
 - b. *breathe down DP's neck* 'put pressure on DP'
 - c. *clip DP's wings* 'restrict DP's freedom'
 - d. *drop DP's name* 'invoke DP for influence'
 - e. *cramp DP's style* 'interfere with DP'
 - f. *get under DP's skin* 'irritate DP'
 - g. *bet DP(self)'s bottom dollar* 'be certain'
 - h. *drown DP(self)'s sorrows* 'drink to relieve depression'

Here, the sequence consisting of possessor and possessed noun functions like a DP in syntax (e.g. *Martin's wings were clipped*, *Frances' style was being cramped*). Hence, a simple transformational analysis, for example in terms of VP-shells, is not an option.

Interestingly, the asymmetry noted for Richards' *give/get* cases persists: though indirect objects, small clause subjects, and possessors are common free positions, second objects, small clause predicates, and possessed nouns are never free positions if other material is idiomatic. For example, there do not seem to be any idioms like *get Croesus' NP* or *find Montezuma's NP*, where NP could be any noun—or if there is, they are exceedingly rare, compared to the very common types mentioned above.

So far all these idioms can be characterized as having open a specifier position: subject, indirect object, small clause subject, or possessor. Specifiers can be parts of idioms as well (recall *The shit hit the fan* etc. in (8) above), but if a specifier is part of an idiom, then so is the lower material; a specifier cannot be part of an idiom, leaving the complement open. Note too that adjuncts may be added to idioms, as in *clip John's political wings* 'restrict John's political freedom' (cf. Nicolas 1995, Nunberg et al. 1994 for some discussion of idiom modification).

The claim that complement positions cannot be 'free' appears to be falsified by such idioms as the following:

- (15)
- a. *keep tabs on DP* 'keep apprised of DP's actions'
 - b. *part company with DP* 'leave DP'
 - c. *close ranks behind DP* 'collectively support DP'
 - d. *close the door on DP* 'rule out DP'
 - e. *fall in love with DP* 'develop intense emotional obsession with DP'
 - f. *hold a candle to DP* 'compare favorably to DP' (NPI)

Each of these examples introduces the free position with a preposition. I will assume that these PPs are merged higher than the idiomatic object in these cases. I will also argue, below, that the specific prepositions involved are not idiomatically listed in the same way as idiomatic nouns and verbs.

2.3 P-complements versus particles. Verb-particle constructions can be analyzed involving a subordinate predicate (the particle) and an open specifier position (the object) (Svenonius 1994, 1996a, 1996b, Ramchand and Svenonius 2002, *inter alia*), in which case there are literally thousands of such idioms. Consider, for example, how opaque the meanings are in (16), a small representative sample.

- (16) Verb-particle constructions
- a. *rip DP off* ‘plagiarize DP’
 - b. *take DP on* ‘challenge DP’
 - c. *clue DP in* ‘inform DP’
 - d. *crack DP up* ‘amuse DP’
 - e. *live DP down* ‘overcome the shame of DP’
 - f. *put DP out* ‘inconvenience DP’

It has been proposed several times that the verb and the particle form a constituent to the exclusion of the direct object at some underlying level of representation, for example by Chomsky (1957) and Johnson (1991) (cf. also Toivonen 2003 for a recent account, focusing on Swedish). On such analyses, the verb and the particle can be separated by movement, but the idiom is not discontinuous in the base representation. Such an approach seems unlikely to extend naturally to all of the important cases, for example it is far from clear how to extend it to the cases of free possessors illustrated in (14) above.

Examining prepositions in idioms, a striking pattern emerges. First, there are many idioms which consist of a preposition plus a noun.

- (17)
- | | | |
|----|----------------------------|--------------------------|
| a. | <i>at will</i> | <i>at a loss</i> |
| b. | <i>at work</i> | <i>under the weather</i> |
| c. | <i>at best</i> | <i>up a creek</i> |
| d. | <i>below par</i> | <i>in a pickle</i> |
| e. | <i>in limbo</i> | <i>on the take</i> |
| f. | <i>in love</i> | <i>on the run</i> |
| g. | <i>in luck</i> | <i>in the hole</i> |
| h. | <i>in gear</i> | <i>for the birds</i> |
| i. | <i>on time</i> | <i>to the letter</i> |
| j. | <i>by hook or by crook</i> | <i>by no means</i> |

This is but a small sample. The prepositions may be somewhat predictable in these cases, but it seems, not entirely. These being constituents, it is not surprising that they can be listed.

The situation can be contrasted with that for verb-preposition combinations. The usual wisdom is that verb-preposition collocations are also frequently idiomatic. Some examples that seem at first to support this view are given in (18).

- (18)
- a. *confide in DP* ‘share a secret with DP’
 - b. *grow on DP* ‘become attractive to DP’
 - c. *get to DP* ‘begin to irritate DP’
 - d. *toy with DP* ‘manipulate DP’
 - e. *guard against DP* ‘be wary of DP’
 - f. *gun for DP* ‘have hostile intentions toward DP’

However, on closer examination it appears that such prepositions are not so idiosyncratic. Although the generalizations are poorly understood, it seems that there are semantic regularities in each of these cases. When a verb appears with a particular preposition, then verbs with similar meanings generally also appear with the same preposition, as illustrated in (19).

- (19)
- | | |
|--------------------------------|--------------------------|
| a. <i>toy with DP</i> | a'. <i>depend on DP</i> |
| b. <i>play with DP</i> | b'. <i>rely on DP</i> |
| c. <i>mess with DP</i> | c'. <i>count on DP</i> |
| d. <i>screw with DP</i> | d'. <i>bank on DP</i> |
| e. <i>fuck with DP</i> | e'. <i>lean on DP</i> |
| f. <i>fool with DP</i> | f'. <i>call on DP</i> |
| g. <i>fight with DP</i> | g'. <i>hinge on DP</i> |
| h. <i>lock horns with DP</i> | h'. <i>base DP on DP</i> |
| i. <i>pick a fight with DP</i> | i'. <i>build on DP</i> |
| j. <i>get tough with DP</i> | j'. <i>rest on DP</i> |

I suggest, then that this is a matter of syntactic selection; just as certain verbs select for particular aspectual features in their complements, verbs may be specified to select for certain prepositions in their complements. This, I suggest, is different from idiomatic listing. To illustrate aspectual selection, consider the following examples.

- (20)
- a. The refrigerator {ceased/stopped} working.
 - b. The refrigerator {ceased/*stopped} to work.
- (21)
- a. The refrigerator {continued/kept} churning out ice cubes.
 - b. The refrigerator {continued/*kept} to churn out ice cubes.

A careful semantic analysis of the differences between *cease* and *stop* and between *continue* and *keep* would presumably explain this distinction. In the meantime, we can descriptively capture such alternations in terms of subcategorization; *keep* subcategorizes for a complement which is [+ING]. Similarly, I suggest that verbs like *depend*, *rely*, *count* and so on subcategorize for a complement which is [+ON], presumably a semantic feature but one which is uniquely expressed in English by the preposition *on*. Thus, *depend on* is not an idiom. The expectation is that the number of such features is not substantially higher than the number of such prepositions; for instance if each of *at*, *by*, *for*, *from*, *into*, *on*, *to*, *of*, and *with* correspond to one to three such features, then there might be around twenty. Compare the remarks of Parsons (1990) regarding English prepositions as expressions of thematic roles.

Such a solution is not possible for idiomatic verb-particle constructions; there are too many subclasses and quirky isolated examples; the situation is better described, in many cases, by listing the particle together with the verb, as an idiom. For example, *give up* means ‘quit,’ but no such meaning can be had from *send up*, *offer up*, *hand up*, or the like. The combination *live down* means something like ‘outlive the notoriety of an event,’ but no such meaning can be had from *survive down*, *reside down*, *last down*, etc. Such examples can be multiplied seemingly indefinitely.

I am not denying that there are fully compositional particle constructions—there are many: cf. completive *up*, or many spatial examples like *throw down/up/over/in/on*, or McIntyre’s (2002) *off* meaning ‘away’ in *fly off*, *sail off*, *send off*, *ship off*, *walk off*, *go off*, *carry off*, etc. What I am suggesting is that verb-particle collocations *can* be idioms, while verb-preposition collocations *cannot*.

This asymmetry, I suggest, between particles and prepositions is part of the more general pattern already observed: roughly, specifier positions can be ‘free’ in idioms, while complement positions cannot be (see Svenonius 2003; 2004a for an analysis of arguments of particles as external arguments of P).

3 Domains for idioms

3.1 Strict separation of the T-domain and the V-domain. One property of idioms which I think has been underappreciated is the separation of the lexical and the functional domains. Verbs regularly form idioms with complements, but almost never with material outside the VP. Thus, there are literally thousands of verb-particle idioms, or verb-noun idioms, but virtually no auxiliary-verb idioms, or adverb-verb idioms. It is not difficult to

imagine an idiom consisting of an auxiliary and a verb; in fact, occasionally they are found, as nouns or nominal modifiers, e.g. *will-call* meaning something like ‘pre-ordered for customer pick-up’ as in *pick your tickets up at the will-call window*; consider also *a can-do attitude*, *a must-have item*, *a Devil-may-care attitude*. None of these sorts of collocations can be used as verbs, nor are there any other auxiliary-plus-main verb idioms.

I claimed above that idioms only have meanings that single words can have; the reverse is also true. Just as there are no idioms which combine a modal and a main verb, there are no single words that have meanings corresponding to a modal plus a main verb. Thus, there is no single verb meaning anything like ‘will like’ or ‘must leave’ or ‘might attack.’ This implies that intensional verbs like *need* have a semantics that distinguishes them from modal structures, i.e. *need* does not mean ‘must have.’

Just as there are no idioms including a main verb and a modal verb, there are no idioms consisting of a main verb plus a tense or (higher) aspect, or extremely few.⁹ Thus, there are no idioms which interact with syntax, but can only appear in some particular tense. This is generally true of verbs as well.¹⁰

To be sure, there are many idioms which include modals and tenses and aspects and the infinitive marker. A few are listed in (22).

- (22)
- a. *need not* ‘don’t have to’
 - b. *have to* ‘must’
 - c. *be going to* ‘in the inceptive stages of’
 - d. *be about to* ‘be immediately before’
 - e. *had better* ‘ought to’
 - f. *be to* ‘be obligated to’ (finite only)
 - g. *have been* ‘have visited/travelled’ (used with locational complement)
 - h. *be DP as it may* ‘despite possibly DP’
 - i. *as DP would have it* ‘as DP claims’

None of these include main verbs (apart from *be* in (22g–h) and *have* in (22i)).¹¹ Thus, I claim that they are T-domain idioms, idioms consisting

⁹*Something’s eating DP* seems to be an exception, as it appears to require progressive aspect. Recall from above that it is also exceptional in containing a subject and a free object position.

¹⁰There are a few odd cases of morphological paradigm gaps, such as *forego*, which many people avoid using in the past tense: *%forewent* (pointed out to me by Andrew Spencer). I assume that the explanation will come from a theory of morphology rather than from a theory of idioms.

¹¹David Pesetsky has pointed out to me the expression *can’t seem to*, which idiomatic-

only of material *outside* VP. Similarly, there are many one-word idioms in the T domain, for example the English modals tend to combine modal and temporal information in an idiomatic way (as can be seen by comparing the more compositional tense-modal-aspect combinations in other Germanic languages).

The last couple of examples in (22) include *as*, which might be a C element. Similarly, the use of *shall* by some speakers of English only in questions can be modelled as an idiom consisting of interrogative C and the modal *shall* (it is incidentally also restricted to first person).

- (23) a. Shall we leave?
b. *We shall leave.

As with idioms containing lexical verbs and nouns, T-domain idioms vary regionally and individually; there are speakers with an idiomatic counterfactual *would have* in (24a), or *might could* ‘might be able to’ in (24b).¹²

- (24) a. %I wish you would have said that earlier.
b. %He might could help you.

Thus, there are clear indications of an important boundary for idioms. There are very many idioms that include material inside VP, but no material outside VP. There are also quite a few idioms that include material from the T-domain, sometimes including C material; these idioms do not include lexical nouns or verbs apart from borderline ‘light’ verbs such as *have* and *be*. Of course, lexical material may be reanalyzed as functional material in the course of becoming a T-domain idiom; historically, this is what has happened with the modals, or in *be going to*.

3.2 Phases. The boundary corresponds to the phase boundary postulated for independent reasons in Chomsky (2000; 2001; 2004a;b). Thus, I take the phase, which is the cyclic unit of lexical access, to set a limit on what can be listed in the lexicon (following Marantz 2001, at least on my reading of that work). Essentially, a verb phrase must be processed for semantic and phonological representations as soon as it can be (‘Impatient Spell-Out,’ Svenonius 2001a;b; 2004b). I have suggested that a verb phrase can be processed (‘transferred,’ in Chomsky’s 2004a terms) when

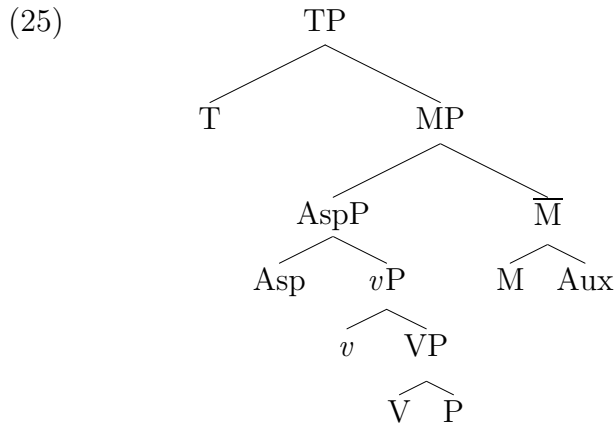
cally means ‘seem to not be able to.’ I would hope that *seem* could be analyzed as a low auxiliary, along the lines of Cinque (2004), rather than as a main verb, at least in this idiom.

¹²Idiomatic for speakers for whom *could* cannot be embedded under any other modal; dialects like the one described in Brown (1991) seem to have more systematic possibilities for combining modals.

all uninterpretable features in it are checked; features are checked as material bearing the relevant features is merged. When all the uninterpretable features are checked the verb phrase, not including T-domain material, is mapped onto semantic and phonological representations and only its label is available to further syntactic operations.

I assume that there is a strict universal functional sequence (f_{seq} of categories, in Starke's 2004 terms). For example, imagine an f_{seq} like $\text{Top} < \text{C} < \text{T} < \text{M} < \text{Asp} < \text{Aux} < v < \text{V} < \text{P}$, such that the output of Merge cannot place a lower member of f_{seq} above a higher one. In this f_{seq} , T could dominate M, but not the other way around. Of course, languages have recursion, so at some point there is a way to place a lower category above a higher one, but I assume that that is a restricted possibility; for example, perhaps only P can dominate a higher category, in this sequence.

Note that nothing requires the f_{seq} to appear as a sequence of complements. For example, M could take Aux as its complement but Asp as its specifier, as in (25).



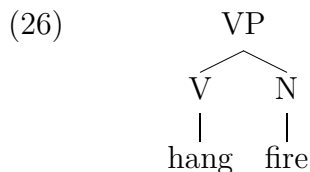
This tree underdetermines the f_{seq} mentioned above, but is compatible with it. It could arise, for example, if the relationship between Aux and V is only indirect.

The general idea behind phases can then be expressed in the following way (based on the version in Svenonius 2004b). Some category, for example V, has a coherent interpretation at the interfaces when all of the features that it dominates are checked (i.e. features of V and features of P, in this parable). Some other categories, for example Asp and v , can check the features that tend to be uninterpretable in V. Thus, what will tend to happen is that when Asp is merged, VP will spell out. VP is therefore called a phase, and Asp and v are called the edge of the phase; material that is attracted to the edge of the phase escapes being spelled out with

the phase. Material outside the phase spells out in the next cycle, as part of a separate phase, according to the same principles: only a node that has coherent interpretations at both interfaces can spell out; this is only true for certain categories, and only when all the uninterpretable features they dominate are checked.

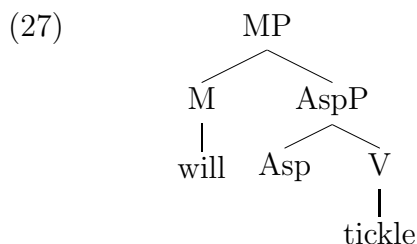
If idioms must be built up by regular rules of syntax before they can be listed in the lexicon, then the phase boundary will interfere with the listing of any unit larger than a phase, as desired.

I will illustrate with a simplified example. Suppose that the idiom is *hang fire*, ‘delay,’ and suppose furthermore that this is stored in the lexicon as a VP with some links to phonological and semantic information, as discussed above.



Suppose that all noun phrases need case, and all verbs have an unbound event variable which must be bound in order for proper interpretation as a property of events; suppose furthermore that these requirements are formally modelled as uninterpretable features. Now, if a functional head *Asp* bearing accusative case feature values and functioning as a binder for the event variable is merged to the VP, then all uninterpretable features are checked and the material in the VP can spell out. At that point it has an interpretation (‘delay’) and a phonological representation (*hang fire*). The label of the constituent is still visible to further operations but its internal structure is not.

Now, suppose that there were an idiom of the form *will tickle* meaning ‘be amusing,’ so that one could say *That movie will tickle* meaning ‘That movie is amusing.’ Such an idiom, by assumption, would include an *Asp* head, since *Asp* by assumption is necessary to bind the event variable present in every verb. The *Asp* head would be located lower than the modal head in the functional sequence, so that the idiom as listed in the lexicon would have to have an entry something like that in (27).



However, by assumption, Asp checks the uninterpretable features of VP, causing it to spell out. Once it is spelled out, its internal properties are not visible to outside operations; it is only a label, linked to phonological and semantic representations. Arguably, this prevents it from being listed together with *will* (or anything else external to VP) as an idiom. That is, the lexicon contains only unspelled-out structures, and a tree like that in (27) cannot exclusively consist of unspelled-out material.

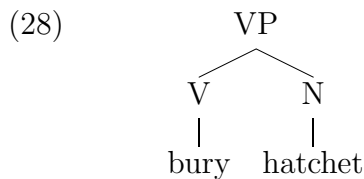
3.3 DP as a phase. The discussion so far seems to suggest that DP cannot be a phase, since DP-internal material frequently forms idioms together with a lexical verb. That might well be correct, and if it is then no major part of the account presented here will have to be changed. Nevertheless, there are some indications that DP is a phase (cf. Svenonius 2004b) and so I want to outline how this is compatible with the account of idioms developed here.

First, though, there are some questions regarding the interpretation of the data. There are many idioms which seem to require definite or indefinite articles (*bite the dust* ‘die,’ *catch a crab* ‘get a cramp’). This seems to suggest that D may be included in an idiom along with N and V. However, it is possible that the choice of *the* or *a* is predictable on the basis of other facts about the idiom, much as the definite article is predictable in expressions like *kick John in the head*, which are not otherwise idiomatic. It is therefore important to examine how many other, less predictable determiners appear in idioms, and they are strikingly rare; there do seem to be a few, including *every which way* ‘every way,’ *in so many words* ‘stated explicitly,’ *leave no stone unturned* ‘look everywhere,’ and there are many with *all*; but I have been unable to find any idioms with *each* or with any of the demonstratives *this*, *that*, *these*, *those*.

Suppose, then, that the noun phrase is structured like the clause, in the way outlined in Svenonius (2004b), with an internal phase that includes the noun, most adjectives, a possessive head, perhaps numerals and the word *all*, but not *each* or demonstratives, which belong higher up, call it the K-domain for case. Suppose furthermore that just as in the clause, the NP phase spells out when its features are checked, and this typically happens

when certain material is merged in the K-domain. The maximal projection of the DP spells out only when it merges in a licensing position outside VP (cf. the Agr_O of Chomsky 1991).

This would mean that, according to the assumptions made so far, N-domain material could be part of an idiom, as long as the relevant part of the K-domain (the edge material that triggers the spelling out of the lower phase) is not present. This is represented schematically in (28) for an idiom like *bury the hatchet*, on the assumption that *the* is introduced later.

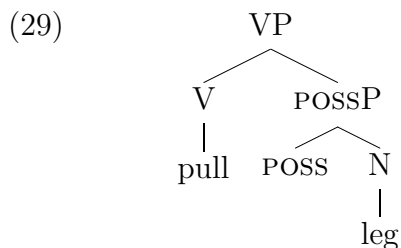


If N contains unchecked features, or if N is not a phase category, then there is no complete phase in the structure here, which therefore can be stored as an idiom. In fact, the problem of inserting the functional structure above N is the same (I argue) as the problem of inserting the specifier material in the various discontinuous idioms discussed above. I turn to this matter in the next section.

What should not be possible is a full noun phrase complement embedded under a verb; a full noun phrase would, by assumption, contain the relevant feature-checkers for the lowest phase in the noun phrase, triggering spell out, and making the listing of the idiom impossible.

4 Merge

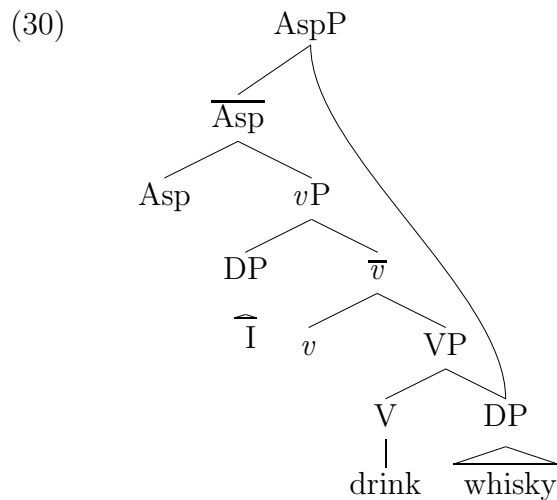
Consider a discontinuous idiom like *pull someone's leg* ‘fool someone.’ One possibility would be to store the structure in (29).



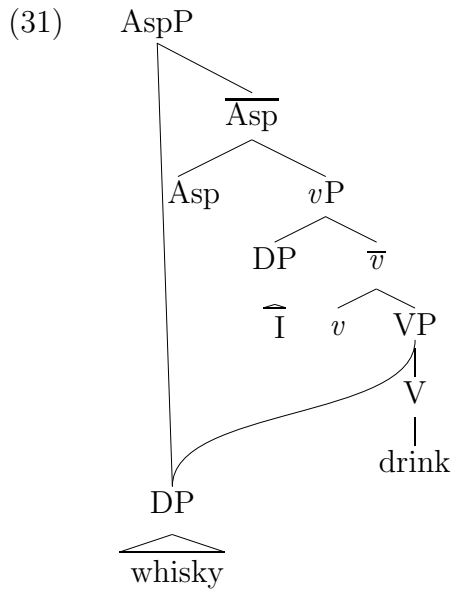
Here, POSS would correspond to the possessive *'s* (cf. Abney 1987) and would require a DP specifier, the ‘free’ position of the idiom. However, merging a specifier to POSS in this tree would require violating the Extension Condition (Chomsky 1993), a basic assumption about the Merge operation, namely that it always “extends its target” in Chomsky’s (1993:22–23) terms.

In this section I show how to resolve the apparent conflict between listing discontinuous idioms as syntactic structures and adhering to the Extension Condition.

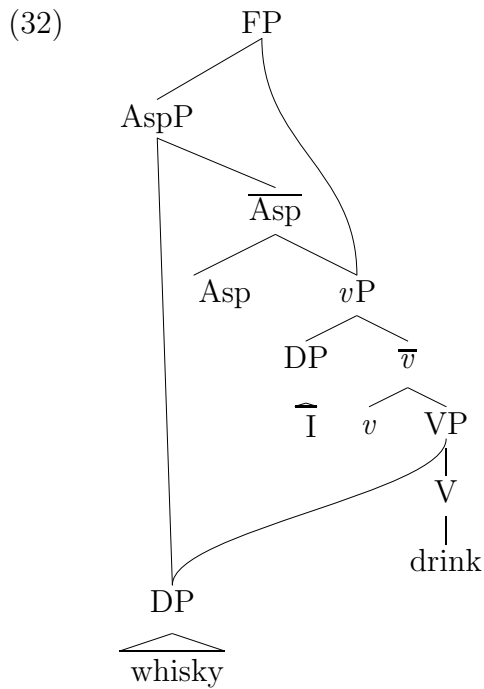
4.1 Merge and Remerge. Merge is standardly assumed to operate on two types of objects: items drawn from the lexicon, and items built by Merge. It takes two such objects and joins them, with one serving as the ‘head’ (cf. Chomsky 1993; 2000). Movement is modeled as the remerging of an already merged node (Starke 2001, Gärtner 2002, Chomsky 2004a), for example if an object DP were to move to SpecAspP for accusative case-checking, that might be diagrammed as in (30), assuming a *v* node introduces the external argument.



There is no copy: the multi-dominated node exists only once. The specifier of AspP is drawn on the right in order to show the complement of V on the right. The diagram could equivalently be represented as in (31).



In (31) the complement of V is drawn on the left, so that the specifier of AspP can be drawn on the left. Since I am not focusing on the details of linearization here, I will generally depict daughter nodes on whichever side is convenient, but it may be generally assumed that specifiers linearize to the left and complements to the right (Kayne 1994). If the licensing movement depicted here is overt, then there must be VP or v P movement to the left of AspP, as illustrated in (32); the head of FP is omitted for simplicity.

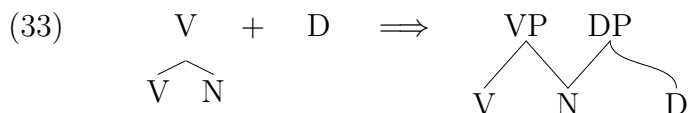


Branches which are not drawn on the ‘correct’ side are represented as curved lines.

Recall the issue raised above regarding DP phases, and the prediction that a complete phasal DP could not be listed under a verb. It may appear, at this point, that the simultaneity of the checking of case on DP and the checking of aspectual features on the verb by Asp is important: it leads to VP and DP spelling out simultaneously. However, this is not crucial. The mechanisms outlined here work even if the edge of the verb phrase is richer, and the case-licenser for DP and the aspectual binder for VP are separate heads, as long as the case-licenser is lower than the aspectual head. What will happen in that case is that the DP will spell out separately from the verb when the case-licenser is merged; for example, the idiomatic object of *hang fire* would receive its idiomatic phonological and semantic representations. When the aspectual head is subsequently merged above the case-licensing head, the verb can spell out as well.

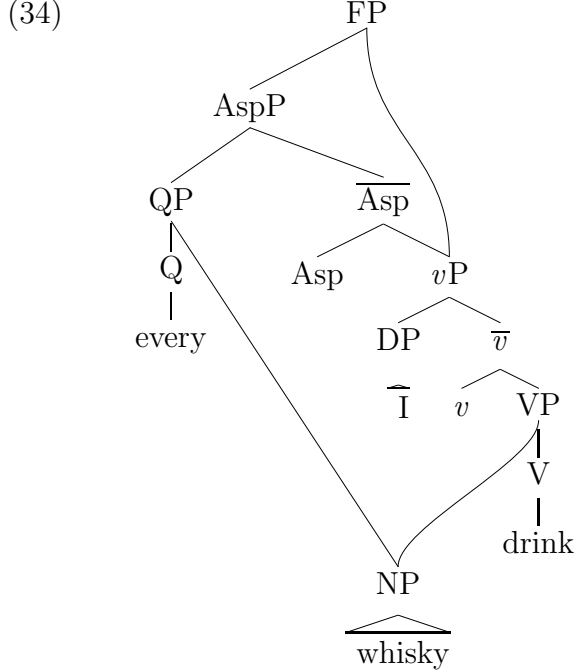
4.2 Banyan trees. I will follow the usual assumption that Merge cannot project a node which is already dominated (this rules out ‘Tucking’ in the sense of Richards 1999; 2001b, though I do not here propose an alternative for the phenomena that Richards discusses). Therefore, Asp in (31) must project when it merges with DP, and F in (32) must project when it merges with *v*P; otherwise, AspP would be adjoined to DP and FP would be adjoined to *v*P instead of the other way around.

However, I will deviate from the usual assumption that only a head merged with an already-constructed structure can target a subpart of it for remerge. In effect, I will assume a limited sort of ‘sideways’ movement (cf. Nunes 2004). For example, suppose V is merged with N. The unusual assumption made here is that D, drawn from the lexicon, can also target the already-merged N. On the assumptions noted earlier that already-dominated material may not project, the only possibility is that D projects, as shown in (33).



This structure is what I call a Banyan tree, after an Indian tree that grows multiple trunks. I suggest that a Banyan tree is not interpretable at the interfaces; the multiple trunks must be resolved under a single trunk by Remerger, before Spell-Out. Assuming that an object DP is checked in SpecAspP, as above, this can be achieved, roughly as sketched in (34); here

I assume the object is a QP, for quantifier phrase.



Assuming all movements here are overt, the verb phrase in this Banyan tree would linearize as *drink every whisky*, as desired (the subject, still containing unchecked features, must move to a higher specifier; see Svenonius 2004b for more details).

For additional arguments that the thematic complement of V is only an N, and not a DP, see Williams (2003) and Sportiche (2005).

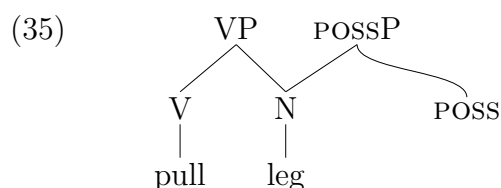
The Banyan tree architecture is fully compatible with conservative assumptions about f_{seq} ; in (34), there is an f_{seq} $F < Asp < v < V$; we can also observe, in the nominal domain, that $Q < N$. If these two sequences are part of the same f_{seq} , then V is higher than N, and Asp is higher than Q, but it is more difficult to determine where Q is in relation to v and V. I would hypothesize that it is higher than both, hence $F < Asp < Q < v < V < N$. I set aside the subject, labeled ‘DP’ here; essentially, if N and V share a single f_{seq} , then the thematic part of the subject must be higher than V, and the case part of the subject must be higher than Asp (see Svenonius 2004b for more detail).

It is expected that the restrictions of f_{seq} and phases will together prevent Banyan trees from overgenerating.

4.3 Idioms and Banyan trees. My claim is that just those idioms are possible that can be represented in Banyan trees that contain no complete

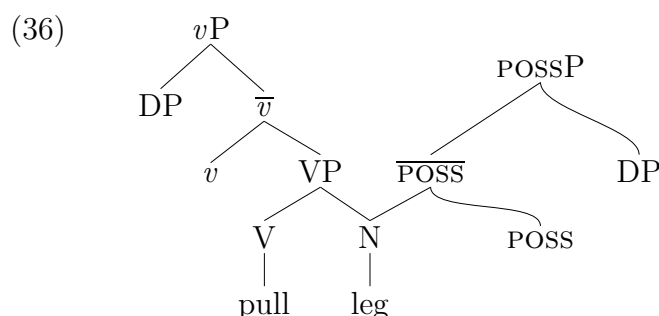
phases. Since the K-domain of DP contains licensers for N, this means there can be no complete DP in an idiom. Since the edge of the verb phrase contains licensers for VP, this means there can be no complete extended verb phrase in an idiom either (hence no auxiliaries, if there is a lexical verb). Otherwise, idioms are free to consist of NP-internal and VP-internal material.

For example, returning now to *pull* someone's *leg*, it can be sketched as in (35).



As before, POSS introduces a specifier DP, but without violating the Extension Condition and without the idiom needing to be listed as separate constituents.

A possessor can be merged with POSSP, and an external argument with VP.



Now, licensing material outside *vP* is responsible for case-licensing the larger object DP, including the possessor. But DP-internal functional heads must be responsible for case-licensing the possessor. Thus, it is possible that the possessor DP spells out before the DP is complete, and the lower (idiomatic) N might as well. At any rate, some higher structure must unify the Banyan tree structure so that both *vP* and POSSP are captured under the same node.

I assume that non-idiomatic constructions with possessive objects are built up in exactly the same way; thus, the thematic object of the verb does not include the possessor, whose relationship to the possessed is calculated separately. That is, if you read Walter's book, there is a thematic reading relationship between you and the book, and there is a possession relation

between Walter and the book, but there is no reading relation between you and Walter. Consequently, the freedom of the interpretation of the possession relation remains regardless of the verb involved: there do not seem to be verbs that restrict the possession relation. Take, for example, *father*; in a sentence like *Joe fathered Mary's child*, the natural interpretation is that Mary was the mother, but other possessive relations are also possible. There does not seem to be a verb that specifies a specific possession relation for a possessor.¹³

4.4 *What can and cannot be part of an idiom.* Now there are some straightforward predictions regarding the interaction of structure and what can be listed. Recall from §2 some of the regularities that were observed for idiom structures.

- (37) Common idiomatic structures
 - a. V–DO: *pass muster* ‘meet with approval’
 - b. V (IO) DO: *read DP the riot act* ‘reprimand DP’
 - c. V (Possessor)’s N: *bend DP’s ear* ‘talk to DP’
 - d. V (DO) Pred: *take DP to the cleaners* ‘cheat DP’
 - e. V (DO) Prt: *rip DP off* ‘cheat DP’
 - f. P–DP: *up a creek* ‘in trouble’
- (38) Uncommon idiomatic structures
 - a. Agent–V: fixed agent subject
 - b. Subject–V (DO): object position open, subject fixed
 - c. V–IO (DO): object position open, indirect object fixed
 - d. V–Possessor (N): possessed open, possessor fixed
 - e. V–DO (Pred): predicate open, object fixed
 - f. V–P (DP): prepositional object open, preposition fixed

The model here can account for these patterns, given some reasonable assumptions about the licensing of DPs in the functional structure. Recall that for each noun phrase, I assume a thematic licenser (e.g. V) and a formal licenser (e.g. Asp). Recall that each licenser has a place in a single, universal, obligatory f_{seq} . Also, when a phase has had all of its features checked, it spells out, and lexical entries cannot contain spelled out material.

First, consider the absence of idioms with agent subjects. There are idioms with subjects (cf. (8) above), but not with agentive subjects. This

¹³Inalienable possession is presumably a property of the possessed noun. For instance, the idiomatic *heart* in *break someone’s heart* is inalienable, hence the possessor must be the heartbroken one.

can be captured if the head thematically introducing an agent is at least as high as the highest head at the edge, that is, is at or above the part of the f_{seq} which causes the verb phrase to spell out (cf. Marantz 1997b). But it would also follow if the head introducing an agent is dominated by the head which triggers the spell-out of VP, e.g. if it were introduced in the specifier of that head.

Second, among idioms with fixed subjects, consider the fact that none has an open object position. This follows from the independently noted fact that thematic subjects are introduced by an element higher than the one introducing thematic objects (Kratzer 1996). If a head v introduces a thematic external argument, then it must embed the lower material, e.g. V; if an object is the complement to V, then it cannot be merged late, but must be part of the idiom or cannot be included at all.

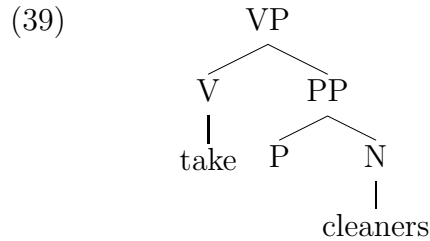
Third, consider the relationship between direct and indirect objects. Indirect objects are apparently not possible parts of idioms. A straightforward explanation would be that the thematic head introducing an indirect object is also the formal licenser for the indirect object (cf. the notion of inherent case in Chomsky 1981). What this would mean is that as soon as an indirect object is merged, it spells out; this would prevent indirect objects from being stored as parts of idioms.

Fourth, a possessor is not typically part of an idiom. The same explanation is available as for indirect objects: if the genitive case on the possessor is assigned by the possessive head, then the possessor DP will spell out as soon as it is merged, and will not be listable as a part of an idiom.

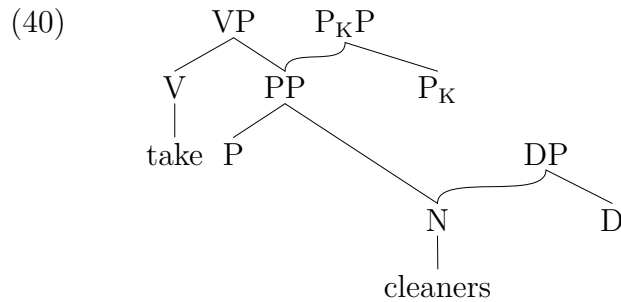
Fifth, a direct object can be free if there is an idiomatic predicate or particle, from which it receives a thematic role, but the predicate cannot be free if there is a fixed direct object. The embedded predicate provides thematic licensing in a specifier position, so the possibility of a free object position is expected. Formal (case) licensing occurs as usual in the aspectual domain at the edge of the verb phrase. But the embedded predicate is a complement of the verb, and therefore cannot be merged late.

Sixth, a preposition can form an idiom with a complement, but not with a selecting verb. The idiomatic complement of P must therefore not be a full, case-licensed DP, because if it were, it would spell out and be unlistable. Thus, the case-assigner in PP must be higher than the lowest projection of P. This suggests that a P which does not assign a thematic role to a noun phrase, but only case-marks it, might be possible as an idiomatic complement to V. But it is not clear that a non-theta-marked noun phrase can be case-marked.

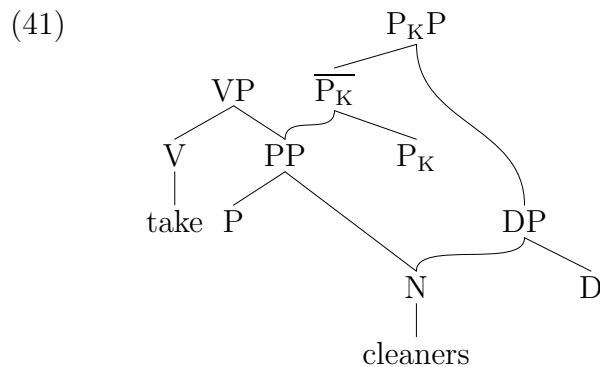
4.5 *An example.* I will illustrate some of these assumptions with the example *take DP to the cleaners* ‘cheat DP,’ mentioned in §1. The idiom itself must be listed as in (39), along with the encyclopedic information that the N represents the state of being stripped of possessions, and so on.



As noted above, there must be a case assigner in PP, call it P_K , which merges to PP; I also assume that all argument noun phrases are DPs, so a D must merge to N.



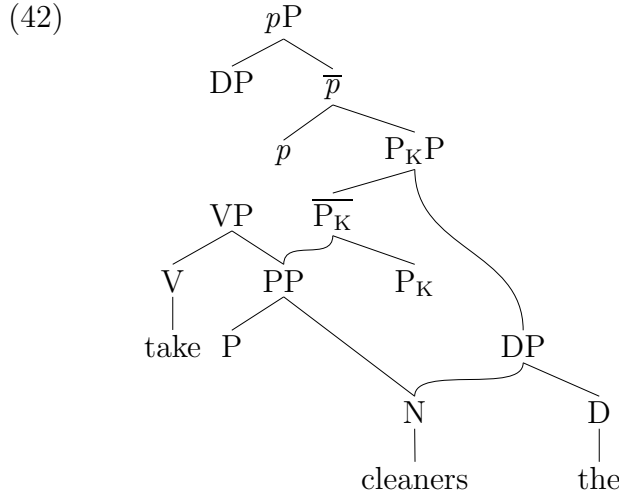
Now, case assignment to DP by P_K could be assumed to happen in a specifier-head configuration (this is not crucial), effected by Merge.



At this point, the DP *the cleaners* can spell out, with its idiosyncratic meaning (‘state of being stripped of possessions’), assuming that a precise motivation can be found for spelling out D as *the*. This P, I assume, is

relational, and requires an external ‘Figure’ argument as well, which is introduced by an additional functional head p (Svenonius 2003); that suggests that spell-out of the preposition does not occur at the stage depicted above.

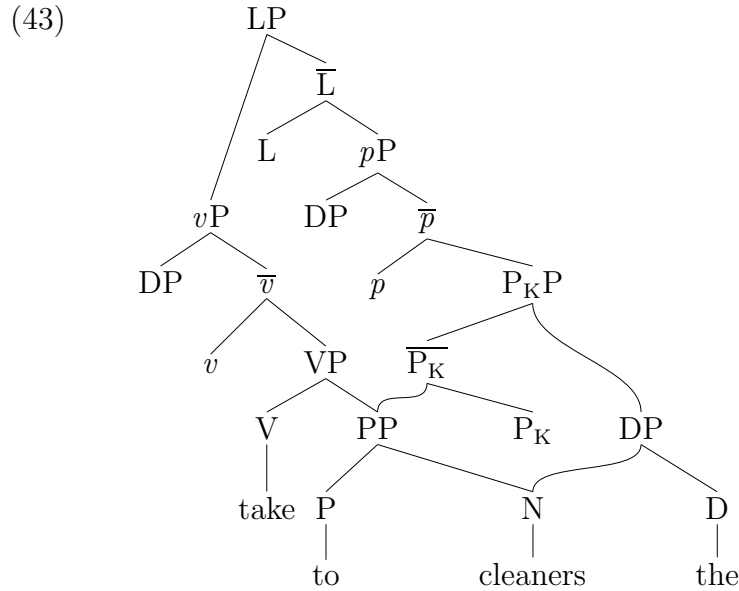
Instead, another head must be merged before the interpretation of P can be made coherent, and this head introduces another argument, the one which will ultimately surface as the direct object.



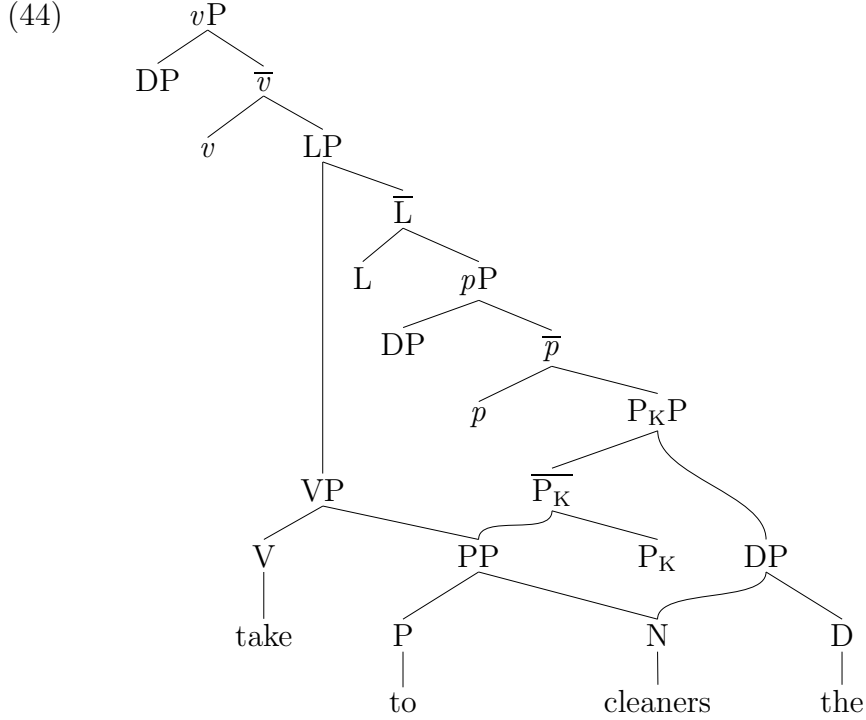
The pP in (42) is a small clause, the situation of the direct object DP going to the proverbial cleaners (it is unclear to me which part of p – P_K – P is lexicalized by *to*, here, but if movement of DP to $\text{Spec}P_KP$ is overt, then it must be p). The VP represents the deliberate process by which this happens. Several more pieces are necessary before this structure can be interpreted. For one thing, an external-argument introducing v must be merged, as must an aspectual head which will check aspectual features in the verb phrase; a case-checking head for the internal argument (the one thematically introduced by p) must also be merged, and it is possibly the very same aspectual head. Finally, something must resolve the Banyan tree into a single root node. Presumably, the relational head which does this also specifies the relation between the process (VP) and the result state (pP). The question, now, is how many different pieces are there and in what order are they merged.

Suppose, for example, that the external argument is introduced by v , and that the relation between the process and result is expressed by a higher head L , for the ‘leads to’ relation corresponding to causation (based on Ramchand’s 2005 ‘leads to’ relation), and that these heads are both lower in f_{seq} than Asp . The Banyan tree, after merge of v and L , would appear as in (43), where the interpretation of L is that the event denoted by vP ‘leads

to' the state of affairs denoted by pP , i.e. the agent's 'taking to cleaners' leads to the theme's 'being at cleaners.'



Merge of Asp would then check the case on the direct object DP and the aspectual features on the verb phrase, allowing both to spell out. However, this structure does not rule out the possibility of an idiom with a fixed subject and an open object position, as long as there is an embedded predicate, like this one. That is because in principle, vP could be listed along with part of the idiom, and the object could be introduced later in Spec pP . This suggests that the tree in (43) uses the wrong f_{seq} , and that the head introducing the external argument must be higher than L, rather than lower.



With the f_{seq} represented in (44), there is no way for an idiom to be listed that contains the external argument but not the internal argument; that is because if the external argument is a fixed part of the idiom, then so is v , and since L is lower, it would have to be part of the structure as well (if there is a secondary predicate), which means that the secondary predicate, which must be the complement of L , must be part of the idiom, which means there cannot be a free specifier position in the secondary predicate. If a verb does not have a secondary predicate, then the object will be a complement of V , which is also lower than v .

These observations motivate an f_{seq} something like $Asp < v < L < p < P_K < P$ and $v < V < P$, which are compatible with each other but underdetermine where p and P_K are with respect to V . Other structures might show whether V can take p or P_K as a subordinate or vice-versa.

5 Reconstruction

One motivation for Banyan trees is the reconstruction facts discussed by Sportiche (2005). The basic observation is that quantifiers frequently fail to reconstruct into the theta positions of the arguments they are associated with. This can be explained if the quantifiers are not merged in the theta positions, but outside them, as I have suggested.

Compare the examples in (45), a pattern originally discussed by Williams (1983).

- (45) a. Somebody seems to be sick.
b. Somebody seems sick.

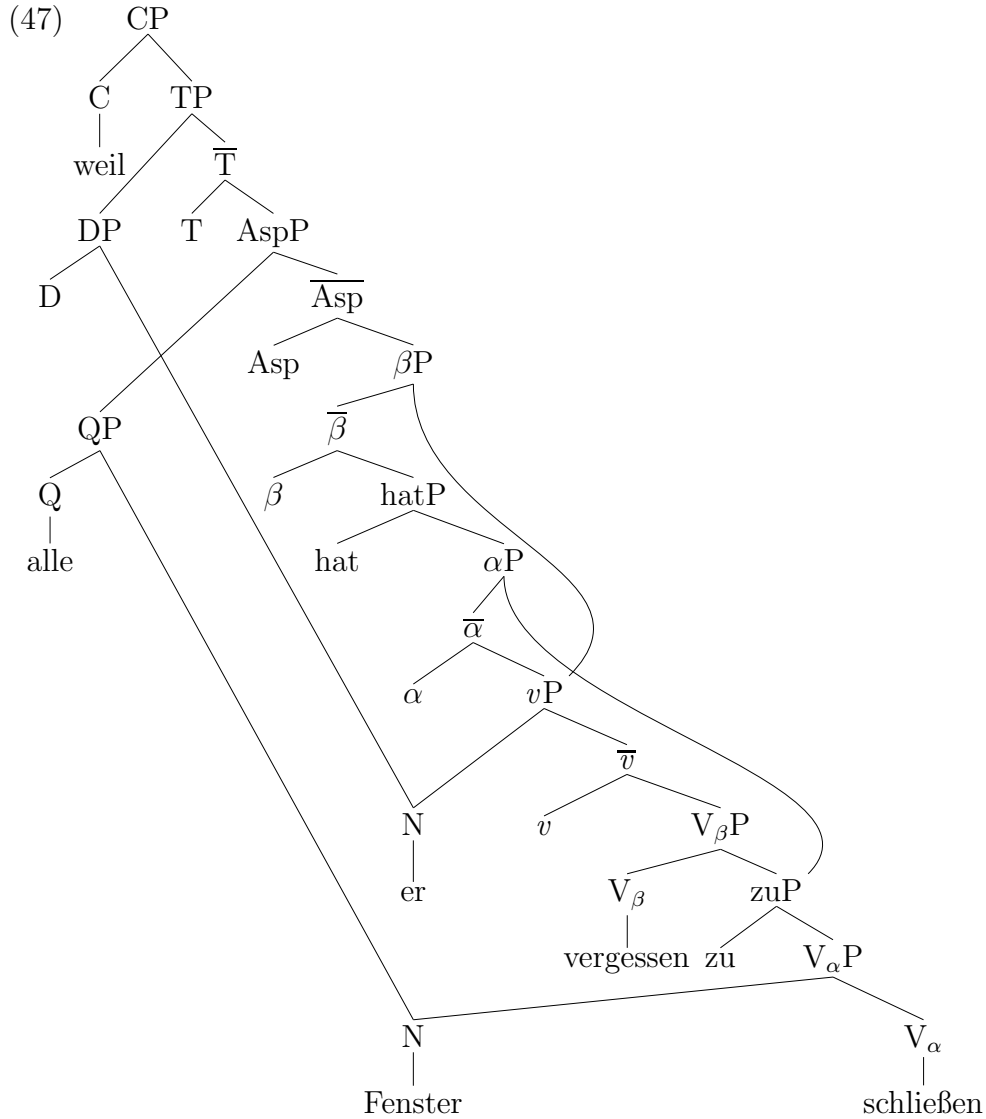
(45a) has two readings while (45b) only has one. On the assumptions here, that suggests that there is a licensing position for *some(body)* in the raising infinitival in (45a), but not in the adjectival small clause in (45b). That is generally plausible; the raising infinitival has a verb, which implies an aspectual head, which was assumed above to provide the right kind of specifier for a QP. An adjectival small clause would lack aspect and hence the licensing position.

Bobaljik and Wurmbrand (2003) point out that reconstruction often fails in restructuring contexts, which Wurmbrand (2001) argues involve VP complements to restructuring verbs. For example, their German example (46), unlike its English translation, only has the reading where ‘all’ outscopes ‘forgot.’

- (46) weil er alle Fenster vergessen hat zu schließen
because he all windows forgotten has to close
‘because he forgot to close all the windows’

On the assumptions made here, a very small complement to a restructuring verb, of the sort Wurmbrand proposes, is expected to lack the licensing positions for quantifiers. The complement of *schließen* is its thematic complement, but the quantifier is introduced higher up. The structure would look something like that in (47), assuming head-initial structures are reversed by nodes α and β . For assumptions regarding the external argument, see Svenonius (2004b).¹⁴

¹⁴Since the verb ‘close’ is resultative, there should be a secondary predicate in (i), which would add some structure but would not change the conclusions here.



6 Conclusion

In this paper I have made two empirical observations about idioms, and proposed a unified solution. The first observation is that the open positions in idioms are very restricted, namely to certain specifiers and adjuncts. Idiomatic material seems to leave complement positions inaccessible.

The second observation is that there is a strict division between the T-domain and the V-domain. There are many idioms in each domain, but very few which span the domain. That is, schematically, the attested kinds of idioms are as in (48b–c).

- (48) a. [TP T [ModalP Modal [VP V [PrtP Prt]]]]
 b. T-Modal idioms
 c. V-Prt idioms
 d. *T-V, *Modal-V, *T-Prt, *Modal-Prt

The solution I proposed is based on phases. Phases provide an independently motivated boundary between the T-domain and the V-domain. Essentially, then, idioms consist of material that fits into a single phase. I suggested that this was because the embedding of one phase inside another requires edge material which necessarily leads to the spell-out of the embedded phase. A spelled-out phase cannot be listed together with material that is external to it.

Given this basic idea about phases, I proposed an analysis of the first observation about idioms, namely that only certain specifier positions remain open in them. Essentially, specifier positions are open if the heads introducing them are not dominated by material. Idioms show that the simplest assumptions about tree structure cannot hold. I introduced Banyan trees, which allow licensing structure to resolve multi-rooted graphs. This allows a restricted form of sideways movement.

References

- Abeillé, Anne. 1995. The flexibility of French idioms: A representation with Lexicalized Tree Adjoining Grammar. In *Idioms: Structural and Psychological Perspectives*, edited by Martin Everaert, Erik-Jan van der Linden, André Schenk, and Rob Schreuder, pp. 15–42. Lawrence Erlbaum, Hillsdale, NJ.
- Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*. Ph.D. thesis, MIT, Cambridge, Ma.
- Bobaljik, Jonathan and Susi Wurmbrand. 2003. Anti-reconstruction effects are anti-reconstruction effects. In *Proceedings of the 2003 Meeting of the Canadian Linguistic Association*, edited by Sophie Burelle and Stanca Somesfalean, pp. 13–24. Université du Québec, Montréal.
- Bresnan, Joan. 1982. Control and complementation. In *The Mental Representation of Grammatical Relations*, edited by Joan Bresnan, pp. 282–390. MIT Press, Cambridge, Ma.
- Brown, Keith. 1991. Double modals in Hawick Scots. In *Dialects of English: Studies in Grammatical Variation*, edited by Peter Trudgill and J.C. Chambers, pp. 74–103. Longman, London.
- Chomsky, Noam. 1957. *Syntactic Structures*. Mouton, 's-Gravenhage.
- Chomsky, Noam. 1980. *Rules and Representations*. Columbia University

- Press, New York.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Studies in Generative Grammar 9. Foris, Dordrecht.
- Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, edited by Robert Freidin, pp. 417–454. MIT Press, Cambridge, Ma.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, edited by Kenneth Hale and Samuel Jay Keyser, pp. 1–52. MIT Press, Cambridge, Ma.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by Step: Minimalist Essays in Honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, pp. 89–155. MIT Press, Cambridge, Ma.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, edited by Michael Kenstowicz, pp. 1–52. MIT Press, Cambridge, Ma.
- Chomsky, Noam. 2004a. Beyond explanatory adequacy. In *Structures and Beyond: The Cartography of Syntactic Structures, vol. 3*, edited by Adriana Belletti, pp. 104–131. Oxford, New York.
- Chomsky, Noam. 2004b. On phases. Ms. MIT.
- Cinque, Guglielmo. 2004. Restructuring and functional structure. In *The Structure of CP and IP*, edited by Luigi Rizzi, pp. 000–000. Oxford University Press, New York.
- Gärtner, Hans-Martin. 2002. *Generalized Transformations and Beyond: Reflections on Minimalist Syntax*. Akademie Verlag, Berlin.
- Hale, Ken and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. No. 39 in Linguistic Inquiry Monograph. MIT Press, Cambridge, Ma.
- Jackendoff, Ray. 1990. *Semantic Structures*. No. 18 in Current Studies in Linguistics. MIT Press, Cambridge, Ma.
- Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. MIT Press, Cambridge, Ma.
- Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9 4: 577–636.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. MIT Press, Cambridge, Ma.
- Kratzer, Angelika. 1996. Severing the external argument from the verb. In *Phrase Structure and the Lexicon*, edited by Johann Rooryck and Laurie

- Zaring, pp. 109–137. Kluwer, Dordrecht.
- Marantz, Alec. 1984. *On the Nature of Grammatical Relations*, vol. 10 of *Linguistic Inquiry Monographs*. MIT Press, Cambridge, Ma.
- Marantz, Alec. 1997a. ‘cat’ as a phrasal idiom: Consequences of late insertion in Distributed Morphology. Ms. MIT.
- Marantz, Alec. 1997b. No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. In *Proceedings of the 21st Annual Penn Linguistics Colloquium*, edited by Alexis Dimitriadis and Laura Siegel, University of Pennsylvania Working Papers in Linguistics, pp. 201–225. University of Pennsylvania, Philadelphia.
- Marantz, Alec. 2001. Words. Ms. MIT.
- McIntyre, Andrew. 2002. Idiosyncrasy in particle verbs. In *Verb-Particle Explorations*, edited by Nicole Dehé, Ray Jackendoff, Andrew McIntyre, and Silke Urban, pp. 95–118. Mouton de Gruyter, Berlin.
- Nicolas, Tim. 1995. Semantics of idiom modification. In *Idioms: Structural and Psychological Perspectives*, edited by Martin Everaert, Erik-Jan van der Linden, André Schenk, and Rob Schreuder, pp. 233–252. Lawrence Erlbaum, Hillsdale, NJ.
- Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70: 491–538.
- Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. MIT Press, Cambridge, Ma.
- Parsons, Terence. 1990. *Events in the Semantics of English: A Study in Subatomic Semantics*. MIT Press, Cambridge, Ma.
- Quang, Phuc Dong. 1971. Sentences without overt grammatical subject. In *Studies Out in Left Field: Defamatory Essays Presented to James D. McCawley*, edited by Arnold M. Zwicky, Peter H. Salus, Robert I. Binnick, and Anthony L. Vanek, pp. 3–10. Linguistic Research, Inc., Edmonton.
- Ramchand, Gillian. 2005. Verb meaning and the lexicon. Ms., University of Tromsø.
- Ramchand, Gillian and Peter Svenonius. 2002. The lexical syntax and lexical semantics of the verb-particle construction. In *Proceedings of WCCFL 21*, edited by Line Mikkelsen and Christopher Potts, pp. 387–400. Cascadia Press, Somerville, Ma.
- Richards, Norvin. 1999. Featural cyclicity and the ordering of multiple specifiers. In *Working Minimalism*, edited by Samuel David Epstein and Norbert Hornstein, pp. 127–158. MIT Press, Cambridge, Ma.
- Richards, Norvin. 2001a. An idiomatic argument for lexical decomposition. *Linguistic Inquiry* 32 1: 183–192.

- Richards, Norvin. 2001b. *Movement in Language: Interactions and Architectures*. Oxford University Press, Oxford.
- Ruwet, Nicolas. 1991. *Syntax and Human Experience*. University of Chicago Press, Chicago. Translated by John Goldsmith.
- Spears, Richard A. 2000. *NTC's American Idioms Dictionary*. NTC, New York, NY.
- Sportiche, Dominique. 2005. Division of labor between Merge and Move: Strict locality of selection and apparent reconstruction paradoxes. Ms. UCLA.
- Starke, Michal. 2001. *Move Reduces to Merge: A Theory of Locality*. Ph.D. thesis, University of Geneva. Available at <http://ling.auf.net/lingBuzz/000002>.
- Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *Structures and Beyond: The Cartography of Syntactic Structures, vol. 3*, edited by Adriana Belletti, pp. 252–268. Oxford, New York.
- Svenonius, Peter. 1994. *Dependent Nexus: Subordinate Predication Structures in English and the Scandinavian Languages*. Ph.D. thesis, University of California at Santa Cruz.
- Svenonius, Peter. 1996a. The optionality of particle shift. *Working Papers in Scandinavian Syntax* 57: 47–75.
- Svenonius, Peter. 1996b. The verb-particle alternation in the Scandinavian languages. Ms. University of Tromsø; available at www.ling.auf.net/lingBuzz/000046.
- Svenonius, Peter. 2001a. Impersonal passives and the EPP: A phase-based analysis. In *Proceedings of the 18th Scandinavian Conference of Linguistics*, edited by Arthur Holmer, Jan-Olof Svantesson, and Åke Viberg, vol. 2, pp. 109–125. Travaux de l'Institut de Linguistique de Lund, Lund.
- Svenonius, Peter. 2001b. On object shift, scrambling, and the PIC. In *A Few from Building E39: Working Papers in Syntax, Semantics, and their Interface*, edited by Elena Guerzoni and Ora Matushansky, MIT Working Papers in Linguistics, pp. 267–289. MITWPL, Cambridge, Ma.
- Svenonius, Peter. 2003. Limits on P: *filling in holes* vs. *falling in holes*. Nordlyd. *Proceedings of the 19th Scandinavian Conference of Linguistics* 31 2: 431–445.
- Svenonius, Peter. 2004a. Adpositions, particles, and the arguments they introduce. Ms. University of Tromsø; to appear in a volume on Argument Structure (*John Benjamins*); available at <http://ling.auf.net/lingBuzz/000042>.
- Svenonius, Peter. 2004b. On the edge. In *Peripheries: Syntactic Edges and*

- their Effects*, edited by David Adger, Cécile de Cat, and George Tsoulas, pp. 261–287. Kluwer, Dordrecht.
- Toivonen, Ida. 2003. *Non-projecting Words: A Case Study of Swedish Particles*. Kluwer, Dordrecht.
- Williams, Edwin. 1983. Against small clauses. *Linguistic Inquiry* 14 2: 287–308.
- Williams, Edwin. 2003. *Representation Theory*. MIT Press, Cambridge, Ma.
- Witkoś, Jacek and Anna Dziemianko. 2005. On the syntax of idioms. Ms. IFA UAM Poznań.
- Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and Clause Structure*. Mouton de Gruyter, Berlin.