

Idioms, Collocations, and Structure

Conventionalized Expressions and the Analysis of Ditransitives

Benjamin Bruening (University of Delaware)

rough draft, January 15, 2018; comments welcome

Abstract

Noncompositional phrasal *idioms* have been used as evidence in syntactic theorizing for decades. A standard assumption, occasionally made explicit (e.g., Larson 2017), is that phrasal idioms differ significantly from fully compositional *collocations* in the kinds of syntactic structures they can be built from. I show with a detailed empirical study that this is false. In fact, the syntactic constraints on idioms and collocations are *identical*. They should therefore be treated the same, as a broad class of *conventionalized expressions*. I show that the selection theory of idioms proposed in Bruening (2010) extends straightforwardly to cover literal collocations. A minor modification accounts for patterns of conventionalized expressions that are striking in their absence. Furthermore, a detailed analysis of the patterns of conventionalized expressions in ditransitive constructions shows that they are systematic and decide between theoretical proposals for the structure of ditransitive constructions in English (contra Larson 2017). Only the ApplP analysis developed in Bruening (2010) accounts for the patterns.

1 Introduction

There is a long tradition in generative syntax of examining the syntactic patterns of non-literal, conventionalized expressions, and using those patterns to inform our theories of syntax, or even to decide between competing analyses of some phenomenon. As an example, patterns of non-literal *phrasal idioms* have been used as arguments for the proper analysis of ditransitive constructions (e.g., Green 1974, Larson 1988, Richards 2001, Harley 2002, Bruening 2010). Larson (2017) dismisses all of these arguments, claiming that expressions that were taken to be phrasal idioms are actually not. Instead they are mere *collocations*, combinations of words that frequently co-occur but do not involve any special, non-compositional meaning. Larson argues that collocations and idioms obey very different structural constraints. According to Larson, there are no true ditransitive idioms, and so idioms are irrelevant for understanding the structure of ditransitives. As for collocations, Larson implies that they are subject to no structural constraints whatsoever, and so are also irrelevant to syntactic theory.

In this paper, I explore whether we should distinguish between different kinds of conventionalized expressions. In particular, I ask whether collocations and idioms truly do obey different structural constraints. A detailed empirical investigation shows that they do not, and in fact are identical. Given this, I further explore what those structural constraints are. I will claim that *selection* is the relevant notion for understanding conventionalized expressions, both literal (collocations) and non-literal (idioms). This is exactly what was proposed for phrasal idioms by Bruening (2010) (building on O’Grady 1998).

Once we understand what the structural constraints on conventionalized expressions are, we can use them to inform our analyses of particular syntactic phenomena. I discuss ditransitives in detail in this paper. Contra Larson (2017), conventionalized expressions are very informative for deciding between syntactic analyses. I show that only the ApplP analysis developed in Bruening (2001, 2010) succeeds in capturing the patterns. All other current analyses, including that of Larson (2014), the HaveP analysis of Harley (2002,

2008) and others, and the transformational analysis of Hallman (2015), are incapable of making the right distinctions.

I begin in section 2 with a discussion of idioms versus collocations and an examination of the syntactic patterns they each appear in. A detailed empirical study shows that they are almost identical in the forms that they take. Section 3 explores the structural constraints that conventionalized expressions are subject to, and shows that a minor modification to the selection theory of Bruening (2010) makes the right distinctions. Section 4 then shows how patterns of conventionalized expressions can be used to decide between analyses of ditransitive constructions in English.

2 Idioms Versus Collocations

Phrasal idioms are typically identified as fixed expressions consisting of multiple morphological (or prosodic) words that, just when they occur together, have a non-compositional interpretation. That is, the phrase as a whole has an interpretation that does not follow from the composition of its parts, interpreted literally. For instance, *kick the bucket* has the non-compositional meaning ‘die,’ and when used in this sense, does not involve any literal kicking or any physical buckets. In contrast, phrasal collocations like *cost a fortune* are interpreted literally (even if somewhat hyperbolically), and their component parts merely co-occur with a high degree of frequency and are recognized as conventionalized expressions by speakers of the language.

In (1) and (2), I list what are typically recognized as differences between idioms and collocations:

- | | |
|--|---|
| (1) kick the bucket (<i>idiom</i>) | (2) cost a fortune (<i>collocation</i>) |
| <ul style="list-style-type: none">• special meaning for combination of particular lexical items• lexical items do not have literal interpretation• non-compositional• lexical items are (generally) non-substitutable (<i>#kick the pail, #boot the bucket</i>) | <ul style="list-style-type: none">• no special meaning, but items frequently co-occur• lexical items have literal interpretation• compositional interpretation• lexical items are substitutable (<i>charge a fortune, cost a bundle</i>) |

Idioms like *kick the bucket* have a special meaning, not their literal meaning; this meaning is non-compositional; and the particular lexical items used cannot generally be replaced with near synonyms. In contrast, collocations like *cost a fortune* are interpreted compositionally and literally, and they often admit substitution of some or all of the lexical items used within them.

Larson (2017) claims that the structural constraints that these two classes of conventionalized expressions are subject to are very different. In particular, he claims that true idioms must be underlying phrasal constituents that exclude all non-idiomatic material, but this is not true of collocations. He does not say what structural constraints collocations are subject to, but implies that there are none. The rest of this section evaluates this claim. A detailed empirical study shows that in fact, idioms and collocations occur in almost identical patterns.

2.1 Constituency

Larson (2017) illustrates his claim that collocations do not have to be underlying constituents with the example of *rancid butter*. According to Larson, this is a collocation. However, it can be freely disrupted by various modifiers:

- (3) rancid yellow creamery butter (Larson 2017: 402 (25))

This is a particularly poor example of a collocation, as *rancid* and *butter* do not actually occur together with any frequency and would not be recognized as a conventionalized expression by any speaker of English. Nevertheless, Larson is correct in that true collocations like *cost a fortune* can be freely disrupted (*cost a small fortune, cost a royal fortune*).

However, this was shown long ago to be possible with phrasal idioms, as well (e.g., Ernst 1981, Nunberg et al. 1994, Nicolas 1995, O’Grady 1998). They also admit adjectival modification:

- (4) a. The Feminist Cat is out of the Marxist Bag (<https://www.youtube.com/watch?v=374dTEMGii0>; idiom: cat out of the bag)
b. First let me cast the usual pall on proceedings! (<https://math.mit.edu/~rbm/papers/anulec/anulec.pdf>; idiom: cast a pall on X)
c. The Bulletin herewith announces its intention to turn over many new leaves. (Bulletin of the Garden Club of America; idiom: turn over a new leaf)

In addition, a very common idiom pattern is for the idiom to consist of a verb and the head of its object, with an open slot for the possessor of the object:¹

- (5) (O’Grady 1998: (4c–e))
a. X lose X’s cool
b. Y get X’s goat
c. Y fill X’s shoes

No theory of phrase structure that I am aware of posits a constituent that includes a verb and the head of its object, while excluding the possessor of the object.

Another common idiom pattern is one where the idiom consists of the verb, an object, and a preposition, but excludes the object of the preposition (O’Grady 1998):

- (6) a. Y beat the bushes for X
b. Y cast a pall on X
c. Y NEG hold a candle to X
d. Y light a fire under X

Again, no theory that I am aware of would posit a constituent that consists of a verb, an NP object, and preposition, but excludes the object of that preposition.

To begin the comparison with collocations, they occur in exactly these same patterns. They allow adjectival modification, as noted above:

- (7) a. hold hands: Do you mind holding sweaty hands?
b. answer the door: Go answer the front door!

They also occur with an open slot for a possessor:

- (8) a. X blow X’s nose
b. Y break X’s heart

¹Some of these idioms require that the possessor of the object be identical to the subject, while others require disjointness. I indicate this with the variables X and Y.

- c. X catch X's breath

And they can consist of a verb, an object, and a preposition, excluding the object of that preposition:

- (9) a. Y break the news to X
- b. Y call/draw attention to X
- c. Y stand trial for X

It is clear that neither collocations nor idioms need to be underlying constituents. Not only that, so far collocations and idioms look almost identical.² To ascertain whether there are any differences between them, I undertook a study of collocations, reported on in the next subsection. As we will see, collocations and idioms are almost identical.

2.2 Empirical Study of Collocations

The empirical study of collocations undertaken here involved the use of the following print collections of collocations:

1. *Oxford Collocations Dictionary for Students of English*. Oxford: Oxford University Press, 2002.
2. Benson, Morton, Evelyn Benson, and Robert Ilson (1997), *The BBI Dictionary of English Word Combinations*. Amsterdam/Philadelphia: John Benjamins.
3. Kjellmer, Göran (1994), *A Dictionary of English Collocations Based on the Brown Corpus*, volumes 1–3. Oxford: Clarendon Press.

The introduction to the Oxford dictionary identifies 13 patterns of collocations. The introduction to the Kjellmer dictionary lists 19 patterns of collocations. The BBI dictionary lists fifteen major types, plus nineteen sub-types involving verbs. All of these are represented in the patterns identified below. Some of the patterns these works view as collocations I treat as simple selection, for instance verbs or adjectives selecting particular prepositions, or adjectives or other categories selecting infinitives (e.g., *ready to*) or *that* clauses. I do not include these here. I also exclude compounds and verb-particle combinations, and include only expressions that are clearly phrasal combinations of multiple words.

I also made use of two on-line lists of collocations:

1. <https://www.worldclasslearning.com/english/collocation-list.html>
2. <https://www.englishclub.com/ref/Collocations/>

The point of this survey is to identify syntactic patterns. For each pattern I identify, I provide representative examples of each. The lists are not meant to be exhaustive. For each pattern that I identify, I also provide a list of phrasal idioms that occur in that pattern (or note that none occur in that pattern). Phrasal idioms were culled from various publications (especially Nunberg et al. 1994, O'Grady 1998, and Bruening et al. 2017), as well as the following idiom dictionaries:

1. *Longman Dictionary of English Idioms*. Harlow and London: Longman, 1979.

²Larson (2017: note 15) also claims that book titles can be non-constituents, citing data from a 2008 *Language Log* post by Geoffrey Pullum (<http://itre.cis.upenn.edu/myl/language-log/archives/005333.html>). This is supposed to show that formulaic language more broadly is subject to no syntactic constraints whatsoever. However, the point of that *Language Log* post is how rare non-constituent titles are and how strained they often feel. I conclude, with Pullum, that these titles show nothing about the syntactic constraints that conventionalized expressions are subject to.

2. Spears, Richard A. and Betty Kirkpatrick (1993), *NTC's English Idioms Dictionary*. Lincolnwood, Illinois: National Textbook Company.

The collocation dictionaries listed above also included idioms.

Note that people may disagree about whether a given phrase is an idiom or a collocation. I took literal meaning of all words in the expression to be the dividing line: if any of the words were non-literal in my estimation, I classified the expression as an idiom. Even then it is not always clear; whether something is non-literal, or just metaphorical or exaggerated, is not easy to determine. In any case, it does not matter, because the patterns are identical, as we will see, and I will argue that there is no reason to distinguish idioms from collocations.

2.2.1 Pattern 1: Modification

The first pattern I identify involves modification, broken down into various subtypes. Pattern 1a has an adjective modifying a noun:

(10) *Collocation Pattern 1a: Modification Adjective-Noun*

active ingredient, adverse effects, big deal, breaking news, bright idea, broken home, clean energy, comfort food, concerted effort, conventional wisdom, crushing defeat, cursory glance, direct quote, distinguishing feature, early days, early riser, easy money, empty promises, false impression, first impression, frayed nerves, free speech, front page, guilty party, happy ending, heavy drinker, heavy losses, heavy smoker, hidden extras, honest mistake, ill effects, informed consent, innocent victim/bystander, large number, little bit, long time, long way, mixed feelings, mysterious circumstances, nasty shock, natural causes, nervous wreck, next time, odd socks, old age, old friend, open wound, painful reminder, plain yogurt, poor health, popular belief, private life, public opinion, quick fix, real life, rhetorical question, ringing endorsement, rough draft, rough estimate, rough idea, round number, runny nose, safe distance, second thoughts, second opinion, sexually transmitted disease, short memory, social life, spare time, stiff competition, straight answer, sure sign, sure thing, tight grip, top priority, top speed, ulterior motive, ultimate goal, unconditional love, unfair advantage, unrequited love, upper echelons, upper limit, used car, vague idea, valid point/reason, valuable information, valuable lesson, vast majority, violent crime, vital organs, vocal critic, warm welcome, warmest regards, weak point, welcome change, wide range, wild animal, wrong number, wrong way, youthful enthusiasm

For comparison, there are also many idiomatic phrases of this sort:

(11) *Idioms: Adjective-Noun*

bad apple, big guns, bitter pill to swallow, busy (little) bee, cold feet, cold fish, cold shoulder, cold turkey, eager beaver, good egg, hot button, hot potato, hot rod, hot water, little bird, the little woman, loose cannon, the old ball and chain, the last straw, X's old man, pretty penny, red herring, shrinking violet, smooth sailing, straight arrow, thin ice, tough row to hoe, white elephant

Pattern 1b has an adverb modifying an adjective:

(12) *Collocation Pattern 1b: Modification Adverb-Adjective*

absolutely necessary, acutely aware, all alone, also known as X, best described as X, completely different, critically acclaimed, crystal clear, dead tired, deeply concerned, deeply held, diametrically opposed, every single, far removed from X, fast asleep, generally accepted, heavily armed, heavily guarded, highly regarded, highly successful, ideally suited (to X), little known, long overdue, painfully shy, painfully slow, perfectly normal, perfectly safe, quite enough, quite good, quite right,

quite sure, readily available, reasonably priced, sound asleep, still alive, strongly opposed, utterly ridiculous, well aware of X, well worth X, wide awake, wide open, widely used

I have only been able to find one phrasal idioms of this sort:

- (13) *Idioms: Modification Adverb-Adjective*
writ large

An alternative analysis of this idiom is that it is an adjectival passive formed from a verb plus secondary predicate, which would put it in pattern 2j below. As for the scarcity of these idioms, it should be noted that idioms are metaphors, and adjectives, being purely descriptive, do not typically take on metaphorical meanings themselves, but only in combination with nouns (as in pattern 1a).

One other thing to note before moving on is that the patterns can combine. Pattern 1a can combine with 1b to produce a collocation *adverb-adjective-noun*, as in *sexually transmitted disease*. I listed this in pattern 1a above; in what follows I will include such complex examples under the heading of just one of the patterns they exemplify, typically the one that I feel is most in need of exemplification (usually because there are not as many examples of that pattern).

Collocation pattern 1c involves modification of a verb by an adverb:

- (14) *Collocation Pattern 1c: Modification Adverb-Verb*
barely able to X, communicate effectively, do better, eat properly, fail miserably, figure prominently in X, get nowhere fast, get off lightly, go ahead, go smoothly, greatly appreciate, know best/better, NEG last long, laugh out loud, love dearly, need badly, never knew (that X), quite agree, recommend X highly, run amok, run rampant, strictly speaking, strongly support X, NEG take long, take X seriously, talk freely, travel light, try hard, turn around, vary widely, word hard, work well

I am only aware of one phrasal idiom with a simple adverb, that in (15a), but there are plenty with adverbial phrases (15b):

- (15) *Idioms: Adverb-Verb*
a. Y only have eyes for X
b. be that as it may, close the stable door after the horse has bolted, X count X's chickens before they hatch, X cross that bridge when X comes to it, Y nip X in the bud, rob Peter to pay Paul, strike while the iron is hot

Collocation pattern 1d involves modification of an adverb or a preposition by another adverb (I treat *just a minute* as an adverb, since it seems to be used as such as a temporal phrase):

- (16) *Collocation Pattern 1d: Modification Adverb-Adverb or P*
hardly any, hardly ever, just a minute, nowhere near, only just, reasonably well, right away, right now, straight after, straight ahead, straight away, way ahead, well after/ahead/before/behind, yet again

I am not aware of any phrasal idioms of this type, but again these would be unlikely to take on metaphorical uses.

Collocation pattern 1e has a noun plus PP. The noun may also be modified by an adjective:

- (17) *Collocation Pattern 1e: (Adjective)-Noun-PP*
article of clothing, bit/piece of advice, blessing in disguise, call of duty, call of nature, cold light of day, enormous amount of X, great deal of X, honor among thieves, kick in the pants, lapse of/in

judgment, momentary lapse of reason, pat on the back, (X's) point of view, prisoner of war, public display of affection, random acts of violence/kindness, ray of light, recipe for disaster, room for improvement, sense of direction, sense of humor, speed of light, thing of beauty, waste of time

There are phrasal idioms of this type:

(18) *Idioms: (Adjective)-Noun-PP*

apple of X's eye, big fish in a small pond, bolt from the blue, boots on the ground, butterflies in the stomach, clean bill of health, dog in the manger, elephant in the room, hair of the dog, land of nod, line in the sand, manna from heaven, pie in the sky, school of hard knocks, shot in the arm, shot in the dark, skin in the game, slap on the wrist, snake in the grass, a snowball's chance in hell, stick in the mud, tall glass of water, tit for tat, wolf in sheep's clothing

Collocation pattern 1f has an adjective followed by a PP (or an NP, in the case of *worth*):

(19) *Collocation Pattern 1f: Adjective-PP/NP*

ahead of time, easy on the eyes, hard of hearing, missing in action, open to suggestions, worth a fortune

There are also at least a few idioms of this form:

(20) *Idioms: Adjective-PP*

blue around the gills, caught in the crossfire, hot under the collar, long in the tooth, running on empty, running on fumes, X too big for X's britches

Collocation pattern 1g involves comparison. These typically involve a comparative adjective plus a standard of comparison, but a few have verbs rather than adjectives:

(21) *Collocation Pattern 1g: Comparison*

as Adj as the next man, better than a kick in the pants, colder than a witch's teat, drink like a fish, dumb as a stump/rock, high as a kite, more dead than alive, old as the hills, pure as the driven snow, quick as a flash/wink, scarcer than hen's teeth, sick as a dog, sleep like a baby/log, slippery as an eel, swear like a trooper, white as a sheet, work like a beaver

There are a couple of idioms of this type:

(22) *Idioms: Comparison*

X's bark is worse than X's bite, may as well be hung for a sheep as a lamb

Comparison serves the same purpose as an idiom: analogizing one thing to something else. It is therefore not surprising that there are few idioms involving comparison (and one could argue about whether the collocations are truly literal).

2.2.2 Pattern 2: Verbs and Arguments/Adjuncts

The second pattern I identify involves a verb and one or more of its arguments, and possibly adjuncts as well (typically PPs). Pattern 2a has a verb and its direct object:

(23) *Collocation Pattern 2a: Verb-Object*

abuse drugs, accept a challenge, accept defeat, answer the door, balance a budget, break a promise, break a record, break the law, catch fire, catch a bus, change the subject, commit a crime, cost a fortune, cover the costs (of X), cut costs, declare war, deliver a baby, do a good/bad job (of X), do

the dishes, earn a living, eke out a living, enter a plea, face a challenge, face facts, find time, follow directions/instructions, get a joke, get the message, give up hope, have a baby, have a big heart, have a chat, have fun, have a good time, have an idea, have no idea (that X), have room, have sex, hold a referendum, hold an election, hold hands, join a club, keep a diary/journal, keep a promise, keep a secret, keep an appointment, keep score, keep the change, keep a low profile, kill time, lead the field, lead the way, leave a message, lose hope, make the bed, make a decision, make a difference, make a fuss, make a living, make a mess, make a mistake, make progress, make a scene, meet a need, meet the requirements, meet a target, miss an opportunity, obey an order, open an account, open fire, pass the time, pass a law, pass a test, pay a bill, pay attention, pull a muscle, put on weight, quit a job, raise money, reach a verdict, reach an agreement, return fire, run a business, satisfy a need, satisfy a requirement, save money/space/time, see reason, serve a purpose, set a date, set a record, set a table, solve a crime, solve a problem, take a call, take a risk, take a seat, take a step, take a test, take ages, take care, take drugs, take liberties, take notes, take notice, take place, take time, take turns, take up space, tell a lie/the truth, tell time, undergo surgery/treatment, waste an opportunity, win an award, yield results

It is well-known that this is also a very common pattern with idioms. Here are a few examples:

(24) *Idioms: Verb-Object*

back the wrong horse, bark up the wrong tree, beat/flog a dead horse, bite the big one, bite the bullet, bite the dust, X bite the hand that feeds X, blow/let off steam, break the ice, bring home the bacon, bury the hatchet, buy the farm, call off the dogs, chew the fat, close up shop, come a cropper, cry wolf, cut corners, dodge a bullet, draw a blank, drive a hard bargain, drop the ball, eat crow, eat humble pie, face the music, fly the coop, foot the bill, gild the lily, NEG give a crap/hoot, go the whole hog, haul ass, have a ball, have an itchy palm, NEG have a leg to stand on, hit home, hit the ceiling/roof, hit the sack/hay, hit the spot, jump the gun, jump ship, kick the bucket, look the other way, lose steam, make tracks, mean business, miss the boat, pack a punch, pass muster, pay the piper, put out fires, rock the boat, ruffle feathers, saw logs, scrape the bottom of the barrel, screw the pooch, shake a leg, shoot the breeze/the bull, sing a different tune, smell a rat, speak volumes, spill the beans, stack the deck, talk turkey, take a powder, throw in the towel, toe the line, turn over a new leaf, turn tail, turn tricks

One thing to note is that the object is almost always inanimate, with both collocations and idioms. Only *baby* and animal names occur. This is yet another way in which idioms and collocations behave alike. (On this tendency with idioms, see Nunberg et al. 1994.)

Collocation pattern 2b was illustrated above (8). It has a verb plus an object, but with an open slot for a possessor of the object:

(25) *Collocation Pattern 2b: Verb-X's Object*

X bear X's teeth, X blow X's nose, Y break X's heart, X catch X's breath, Y catch X's eye, Y catch X's attention, Y come to X's rescue, Y cross X's mind, X do X's best, X do X's duty, X get X's kicks from Y, X give (X's) consent, Y hurt X's feelings, X kick up X's heels, Y lose X's place, X lose X's temper, X lose X's way, X make up X's mind, X pay X's respects (to), X pick X's nose, X put up X's hand, X raise X's voice, Y save X's life, X save X's strength, Y take X's advice, Y take X's temperature, Y take X's place, X take X's (sweet) time, X thank X's lucky stars, Y upset X's stomach, X wait X's turn, X watch X's weight

Phrasal idioms of this type were also illustrated above (5). Here is a longer list of such idioms:

(26) *Idioms: Verb-X's Object*

Y bend X's ear, X bet X's bottom dollar, X blow X's stack, Y breathe down X's neck, X burn X's bridges, Y clip X's wings, Y cook X's goose, Y cramp X's style, Y fill X's shoes, Y get X's goat, X hold X's tongue, X let X's hair down, X lose X's cool, X lose X's head, X keep X's shirt on, Y pull X's leg, X put X's foot down, X spill X's guts, Y steal X's thunder, X tighten X's belt

Collocation pattern 2c was also illustrated above (9). It involves a verb, a direct object, and a preposition, with an open slot for the object of the preposition (note that the subject of all of them is disjoint from the open object of the preposition):

(27) *Collocation Pattern 2c: Verb Object P X*

bear/hold a grudge against X, bear a resemblance to X, break the news to X, call/draw attention to X, cast doubt on X, catch a glimpse of X, catch sight of X, come to terms with X, do damage to X, draw the line at X, express interest in X, gain/have access to X, get a bang out of X, get a kick out of X, get rid of X, give birth to X, give chase to X, give consideration to X, give credence to X, give form to X, give priority to X, give some thought to X, give way to X, have a word with X, have an effect on X, impose conditions/restrictions on X, join forces with X, keep an eye on X, lay the blame on X, lay the groundwork for X, look forward to X, lose control of X, lose faith in X, make a note of X, make arrangements for X, make friends with X, make short work out of X, open fire on X, part company with X, pay attention to X, pay homage to X, pay respect(s) to X, pay tribute to X, place an order for X, pose a risk to X, pose a threat to X, put pressure on X, put a stop to X, raise doubts about X, raise questions about X, regain control of X, run the risk of X, say goodbye/hello to X, say sorry to X, set fire to X, Y set Y's sights on X, stand trial for X, strike a balance between X, take a break from X, take a look at X, take action against X, take an interest in X, take care of X, take charge of X, take issue with X, take part in X, take pride in X, tell the difference between X, wage war against X, wave goodbye to X

Idioms of this pattern were also illustrated above (6). A longer list of such idioms appears below. Note again that the subject is always disjoint from the object of the P (yet another way in which idioms and collocations are identical):

(28) *Idioms: Verb-Object P X*

beat the bushes for X, carry a torch for X, cast a pall on X, do a number on X, get the hang of X, get a kick out of X, get a load of X, give it to X, give rein to X, give rise to X, give the lie to X, give vent to X, give voice to X, X hang X's hat on Y, have a bone to pick with X, have an axe to grind with X, NEG hold a candle to X, keep a lid on X, keep tabs on X, light a fire under X, lock horns with X, make head or tail of X, pay court to X, pay lip service to X, pour cold water on X, pull the plug on X, Y put Y's finger on X, shed/cast light on X, Y sink Y's teeth into X, take a back seat to X, take a gander at X, take a shine to X, think the world of X, throw the book at X, turn a blind eye to X

Collocation pattern 2d involves ditransitives. Most of them are listed in my sources as the double object variant. These include a verb and second object, but with an open slot for the first object. I list them in two groups below, depending on whether or not they alternate, in my judgment. Those that alternate alternate with a prepositional variant, for instance, *do a favor for X* or *give advice to X*. (Individual speakers may vary in which they allow to alternate.)

(29) *Collocation Pattern 2d: Verb X NP, Non-Alternating*

call X names, deny X access to, do X good/harm, give X permission, give X a bad name, give X a break, give X a call, give X a chance, give X a go, give X a lift, give X a minute, give X a round of applause, Y give X Y's word, give X some idea of Y, give X the benefit of the doubt, give X the impression that Y, save X the trouble, tell X the time

- (30) *Collocation Pattern 2d: Verb X NP, Alternating*
do X a favor, give X advice, give X credit, give X meaning, give X some guidance, give X some thought, give X short shrift, lend/loan X money, make X an offer, offer X a job, offer X an explanation, offer X (Y's) condolences, offer X Y's services, owe X an apology, owe X an explanation, pay X a compliment, pay X a visit, save X a seat, teach X a lesson, tell X a thing or two

Idioms also occur in this pattern. Larson (2017) denies that these exist, but he only discusses those in (31). There are numerous others, some of which are listed in (32):

- (31) *Idioms: V X NP*
give X the boot, give X the creeps, give X flak, give X a kick, Y give X a piece of Y's mind, promise X the moon, read X the riot act, show X the door, show X the ropes
- (32) *Idioms: V X NP*
Y bet X Y's bottom dollar, bet X dollars to donuts, cut X some slack, cry X a river, drop X a line, give X a free hand, give X a hand ('applaud' or 'help'), give X pause, Y give X the rough edge of Y's tongue, give X a turn, give X a whirl, give X no quarter, give X the once-over, give X the slip, give X a wide berth, give X what-for, give X some skin, lend X color, Y lend X Y's ear, lend X a hand, NEG pay X any mind, throw X a bone, throw X a curve

Collocation pattern 2e consists of a small number of double object collocations where both NPs are part of the expression:

- (33) *Collocation Pattern 2e: V NP NP*
give it a rest, give it your best shot, give it the old college try, don't give me that (Imperative)

There are at least two idioms of this form, as well:

- (34) *Idioms: V NP NP*
give the devil his due, give X's self airs

The dividing line between literal collocations and non-literal idioms is not so clear here. I will return to conventionalized expressions involving ditransitives in section 4.

Collocation pattern 2f has a verb, a direct object, and a PP, all fixed (but possibly with an open possessor):

- (35) *Collocation Pattern 2f: V NP PP*
Y bring a tear to X's eye, X get a lump in X's throat, X give a good account of X's self, take the rough with the smooth

There are many more idioms of this sort:

- (36) *Idioms: V NP PP*
add fuel to the fire, build castles in the air, burn the candle at both ends, cast pearls before swine, change horses in midstream, get a foot in the door, get the show on the road, give hostages to fortune, go to hell in a handbasket, have a bun in the oven, hit the nail on the head, Y hold X's feet to the fire, keep the wolf from the door, kill two birds with one stone, X lay X's cards on the table, X paint X's self into a corner, X place X's head in the lion's mouth, Y pull the wool over X's eyes, X put all X's ducks in a row, put the cart before the horse, take/seize/grab the bull by the horns, Y take the wind out of X's sails, Y take the word's out of X's mouth, throw out the baby with the bath water

Collocation pattern 2g has just a verb and a PP:

(37) *Collocation Pattern 2g: Verb PP*

come into view, come to a stop, come to a standstill, come to an agreement, come to an end, fall in love, get into trouble, go out of fashion, fly by night, jump at shadows, jump to conclusions, look for trouble, meet with approval, resort to violence

There are also many idioms of this form:

(38) *Idioms: Verb PP*

bark up the wrong tree, beat around the bush, cut to the chase, dance with the devil, fall in(to) line, fly off the handle, fly under the radar, get in on the ground floor, get into the swing of things, get to first base, Y get under X's skin, go overboard, fall in line, is in the wind, jump on the bandwagon, knock on wood, pay through the nose, play with the big boys, Y rain on X's parade, read between the lines, X sit on X's hands, Y slip through X's fingers, start from scratch, X stew in X's own juices, X stick to X's guns, swim against the current, swim with the sharks, X take to X's heels

Collocation pattern 2h has a verb and a PP, but an open slot for an object in between (in all of these collocations, the PP appears to be an argument of the verb):

(39) *Collocation Pattern 2h: Verb X PP*

bring X to an end, bring X to justice, drive X to distraction, keep X in mind, keep X under wraps, put X out of X's misery, take X to court, Y tell X to X's face

There are also many idioms of this form (with both argument and adjunct PPs):

(40) *Idioms: Verb X PP*

carry X to extremes, can cut X with a knife, drag X over the coals, feed X to the lions, hoist X with X's own petard, Y keep X under Y's hat, kick X to the curb, lay X on the line, Z lay X at Y's door, leave X in the lurch, push X to the edge, put X out to pasture, rake X over the coals, sell X down the river, send X to the devil, send X to the showers, stab X in the back, sweep X under the rug, take X to the cleaners, take X to task, Y take X under Y's wing, take X with a grain of salt, throw X to the wolves, throw X under the bus

The subject is always disjoint from the object, in both collocations and idioms.

Collocation pattern 2i has two PPs, with an open slot for the object of the second:

(41) *Collocation Pattern 2i: Verb PP P X*

come to grips with X, go on a date with X, keep in touch with X

I have not found any idioms of this form.

The following two cases do not fit with any others, so I will call them both Pattern 2j. In the first, there is a verb and PP, with an open slot inside the object of the P. In the second, which typically occurs with negation, there is simply a verb and PP with an open slot for the object of the P:

(42) *Collocation Pattern 2j: Verb P... X*

fall into the category of X, NEG bet on X,

It would be surprising if there were no others of this form. I have also found no idioms of either type.

The next two patterns involve verbs and adjectives. Collocation Pattern 2k has a verb plus adjective, where the adjective is either selected by the verb or is a secondary predicate modifying the subject of the verb:

- (43) *Collocation Pattern 2k: Verb Adjective/Secondary Predicate*
 be dragged kicking and screaming, become apparent (that X), call in sick, come alive/ to life, come close, come in first/last, come true, drop dead, fall asleep, get ready, get upset, get worried, go bald, go bankrupt, go crazy, go unchallenged, go unnoticed, go wrong, keep busy, keep/stay calm, keep fit, keep mum, keep quiet, keep still, keep warm, look nice, make sense, make sure, run late, run low, sit crosslegged, stay awake, stay put, stay tuned
- (44) get something straight, get/set the record straight, X keep X's ears/eyes open, X keep X's mouth shut, make a long story short, take it easy

I also include here as a sub-case some collocations that also include a direct object (44). There are some idioms consisting of a verb and adjective or secondary predicate:

- (45) *Idioms: Verb Adjective/Secondary Predicate*
 getting busy, sitting pretty, play hard to get
- (46) X get X's feet wet, get/keep/start the ball rolling, Y got X's panties in a bunch, have it made, keep the home fires burning, X keep X's shirt on (Imperative), NEG leave stone unturned, let sleeping dogs lie, paint the town red, X sell X's self short

Some also include an object (some with an open slot for a possessor). I include *X keep X's shirt on* here although it has a particle rather than an adjective, because the order particle-NP is not possible.

Collocation pattern 2l has a verb and adjective or secondary predicate, but also has an open slot for a direct object (the subject of the secondary predicate, necessarily disjoint from the subject of the verb):

- (47) *Collocation Pattern 2l: Verb X Adjective/Secondary Predicate*
 beat X to death, bring X (a)round, drive X crazy, get X ready, have X removed, hold X hostage, hold X prisoner, keep X waiting, keep X quiet, leave X alone, make X better, make X easy (for), make X public, put X together, read X aloud/out loud, serve X right, set X on fire, set X free, take X apart, take X hostage/prisoner, tell X apart

There are few idioms of this form; I have found the following:

- (48) *Idioms: Verb X Adjective/Secondary Predicate*
 give it to X straight, keep X posted, knock 'em/X dead, pull it/X together

The first one has a fixed object but open object of P.

Collocation pattern 2m has a verb plus an infinitive, sometimes with an open slot for an object or with a fixed object:

- (49) *Collocation Pattern 2m: Verb (NP/X) Infinitive*
 give X to understand Y, X have it in X to Y, have the nerve to X, lead X to believe Y, let (X) go, let X know, quit drinking/smoking, talk X into/out of Y

There are three idioms that I have found of this form:

- (50) *Idioms: Verb (NP/X) Infinitive*
 let sleeping dogs lie, let them/X eat cake, make X's blood boil

Finally, there is a small number of collocations that include a verb and a CP argument. One of these has an NP argument in addition to a CP argument:

- (51) *Collocation Pattern 2n: V (NP) CP*
X is as X does, say/tell me it isn't/ain't so, tell/know which is which

In *X is as X does*, the clause with *as* seems to be the complement of the verb *be*. It is a full finite clause. The subject of the lower clause is obligatorily covalued with the matrix subject (and typically repeats it, e.g., *famous is as famous does*). There are a few idioms that have CP arguments, possibly in addition to an NP argument:

- (52) *Idioms: V (NP) CP*
know what the score is (*or* know the score), look what the cat dragged in, X teach X's grandmother to suck eggs, tell X where to get off

Know what the score is also occurs as *know the score*, with just an NP object (but that NP object is semantically a concealed question).

This concludes the patterns of collocations that I have found involving a verb and its arguments and adjuncts.

2.2.3 Other Patterns

The rest of the collocation patterns I have identified are of various different forms. Collocation pattern 3 involves coordinations:

- (53) *Collocation Pattern 3: Coordinations*
- a. aches and pains, bits and pieces, (this) day and age, everybody and his uncle, for all intents and purposes, health and wellbeing, law and order, nooks and crannies, nothing but net, peace and quiet, time and time again, this and that, thunder and lightning, wear and tear
 - b. bait and switch, cut and paste, give and take, grin and bear it, kiss and tell, stop and go (traffic), suck it up and deal, take it or leave it (Imperative), tax and spend, toss and turn
 - c. alive and kicking, black and blue, black and white, bound hand and foot, more or less, neat and tidy, null and void, safe and sound
 - d. down and out, over and above
 - e. and yet, but again, or else

There are also idioms that are coordinations:

- (54) *Idioms: Coordinations*
betwixt and between, the birds and the bees, bow and scrape, by and large, by hook or by crook, fire and brimstone, (NEG) make head or tail of X, the old ball and chain, one foot in and one foot out, pins and needles, shit or get off the pot, (on) the up and up

Collocation pattern 4 has a preposition followed by its NP object:

- (55) *Collocation Pattern 4: Preposition NP*
above all, above/under ground, against all odds, against the law, after all, along those/the same lines, at first sight, at some point, beside the point, between jobs, beyond recognition, beyond reproach, by accident, in advance, in agony, in any case, in black and white, in short, of a personal nature, on (the) alert, on loan (from X), out of order ('not functioning' or 'out of sequence'), over the edge, under glass, under the microscope, under the circumstances

There are idioms of this form, as well:

(56) *Idioms: Preposition NP*

after a fashion, at sixes and sevens, behind the eight ball, below the belt, beside X's self, between the devil and the deep blue sea, between a rock and a hard place, beyond the pale, down in the dumps, down that road, in the black/red, in the can, in the doghouse, in the same boat, off the grid, on the hop, on the double, on the dole, out of line, out on a limb, over a barrel, over X's head, over the hill, over the hump, under the weather, under wraps

Collocation pattern 5 has an adjective plus infinitive:

(57) *Collocation Pattern 5: Adjective Infinitive*

too good to be true, hard to believe, hard to say, hard to take, hard to tell, needless to say, too hot to handle

There is at least one idiom of this form:

(58) *Idioms: Adjective Infinitive*

fit to be tied

Collocation pattern 6 is predicative. It has an NP that is the subject of a small clause:

(59) *Collocation Pattern 6: Predicative (NP + Small Clause)*

hand in hand, hands off, head over heels, no money down, no strings attached, time off, upside down

There are some idioms of this form as well:

(60) *Idioms: Predicative (NP + Small Clause)*

back (is) against the wall, blind leading the blind, cat (is) out of the bag, fat (is) in the fire, noses (are) to the grindstone, tail (is) wagging the dog, writing (is) on the wall

Collocation pattern 7 involves one or more functional elements plus a main verb and an open slot, possibly with other elements (like the object of the verb):

(61) *Collocation Pattern 7: F V (NP P) X*

- a. has had enough of X, had better X, would rather X, to say nothing of X
- b. can't afford X, can't do any harm, can't help X-ing, can't stand X, couldn't/don't give X (where X is a minimizer)

I list as a separate subtype collocations that have specifically negation, a modal, and a verb (61b), as there are several of these. There are some idioms that include functional elements, some have open slots while others do not:

(62) *Idioms: F V (NP P X)*

be cruisin' for a bruisin' (progressive), break a leg! (imperative), can cut X with a knife, can/could NEG help it/Xself ('unable to stop oneself'), has been around the block a few times, has had it, has gone to the dogs, no can do

See section 3.1 on functional elements outside the VP, and also section 3.2.3 on negation as a functional element in many idioms.

Collocation pattern 8 has a verb, an object, and an open clausal constituent. In (63a), this is probably a complement of the verb, while in (63b), it is probably the complement of the noun that is the object:

(63) *Collocation Pattern 8: V (X's) NP Y*

- a. X spend X's life Y-ing, have a hard time X-ing
- b. get the impression that X

I have found one idiom of this form:

- (64) *Idioms: V (X's) NP Y*
X cut X's teeth Y-ing

Collocation pattern 9 involves a full clause, including a subject:

- (65) *Collocation Pattern 9: Full Clause*
- a. duty calls, easy does it, nature calls, the prodigal son returns, X's heart is in the right place, what you see is what you get, you know
 - b. What/something is up, What gives?, What's up with X?, How are you?, Who's there?
 - c. when all is said and done, the fact remains that X, given time (adjunct clause with null controlled subject)

There are also idioms that are full clauses, some with open slots for possessors or objects of prepositions:

- (66) *Idioms: Full Clauses*
- a. the fur fly, that ship has sailed, the die is cast, the jig is up
 - b. the shit hit the fan, Bob's your uncle, loose lips sink ships, something is on the wind, what/something crawled up X's ass and died
 - c. what/something is eating X
 - d. until the cows come home, when hell freezes over, come hell or high water
 - e. the bottom fell out of X, the roof caved in on X

Finally, collocation pattern 10 consists of two clauses in juxtaposition, with a conditional-like meaning (note that the second clause can have a *then* added, just like a conditional: *no pain, then no gain*). I will call these *covert conditionals*:

- (67) *Collocation Pattern 10: Covert Conditionals*
easy come, easy go; no pain, no gain;

There is at least one idiom of this type:

- (68) *Idioms: Covert Conditionals*
monkey see, monkey do

2.3 Summary of Empirical Study

I summarize the patterns identified here below, with an example of a collocation of that form and a matching idiom if there is one. The sub-types of Pattern 1 involve modification:

- (69) *Pattern 1: Modification*
- a. Adj N: honest mistake (idiom: cold feet)
 - b. Adv Adj: acutely aware (idiom: writ large)
 - c. Adv V: go smoothly (idiom: only have eyes for X)
 - d. Adv Adv/P: hardly ever, way ahead (no idioms)

- e. N PP: call of duty (idiom: elephant in the room)
- f. Adj PP/NP: missing in action, worth a fortune (idiom: running on fumes)
- g. Comparison: better than a kick in the pants (idiom: X's bark is worse than X's bite)

The sub-types of Pattern 2 involve a verb and one or more of its arguments, or adjuncts to the VP:

(70) Pattern 2: Verbs and Arguments/Adjuncts

- a. V NP: answer the door (idiom: back the wrong horse)
- b. V X's NP: Y break X's heart (idiom: Y clip X's wings)
- c. V NP P X: break the news to X (idiom: beat the bushes for X)
- d. V X NP: give X permission, do X a favor (idiom: cut X some slack)
- e. V NP NP: give it a rest (idiom: give the devil his due)
- f. V NP PP: Y bring a tear to X's eye (idiom: build castles in the air)
- g. V PP: come into view (idiom: beat around the bush)
- h. V X PP: bring X to justice (idiom: feed X to the lions)
- i. V PP P X: keep in touch with X (no idioms)
- j. V P... X: NEG bet on X (no idioms)
- k. V (NP) Adj: call in sick, make a long story short (idiom: getting busy, X get X's feet wet)
- l. V X Adj: drive X crazy (idiom: pull it/X together)
- m. V (NP/X) Infinitive: lead X to believe (idiom: let sleeping dogs lie)
- n. V (NP) CP: tell me it isn't/ain't so (idiom: X teach X's grandmother to suck eggs)

Patterns 3–10 involve other structures:

- (71) (P3) Coordinations: law and order (idiom: the old ball and chain)
- (72) (P4) P NP: against the law (idiom: behind the eight ball)
- (73) (P5) Adj Infinitive: hard to say (idiom: fit to be tied)
- (74) (P6) Predicative: no money down (idiom: back against the wall)
- (75) (P7) F V (NP P) X: can't stand X (idiom: has had it)
- (76) (P8) V (X's) NP Y: X spend X's life Y-ing (idiom: X cut X's teeth Y-ing)
- (77) (P9) Full Clause: easy does it (idiom: the shit hit the fan)
- (78) (P10) Covert Conditional: easy come, easy go (idiom: monkey see, monkey do)

As can be seen, collocations and idioms overlap significantly. The only collocation patterns where there does not appear to be a corresponding idiom of that form are Patterns 1d, 2i, 2j. Some of these patterns have very few examples of collocations (2i, 2j), so it may not be surprising that there are no idioms of that form. Pattern 2i is a little surprising both because of how few collocations of that form there are, and in that there are no idioms of that form. My guess is that there are examples of both that I have simply not found yet. One P NP idiom, *out on a limb*, frequently occurs as *go out on a limb for X*, which fits this pattern, but it also occurs without the verb and without the *for* PP. As for pattern 1d, it is headed by an adverb, and it is unlikely that adverbs would take on metaphorical meanings and so become idioms.

The overall correspondence between collocations and idioms is striking. There are few cases of non-overlap, and those either have a good explanation (adverbs not being used metaphorically) or are probably not significant (low numbers for the one that does occur). This indicates that any constraints on the structure

of idioms are shared by collocations, and vice versa. That is, they belong to the same category, which I will call *conventionalized expressions*. Importantly, there is no support for Larson's 2017 contention that collocations and idioms are subject to very different structural constraints. Instead, we want a unified theory of both idioms and collocations.

A few more observations are in order concerning the data presented in this section. First, thematic roles of all kinds occur in both collocations and idioms. This is especially true of PPs and will be relevant in section 4. There are PPs that encode goals (*jump on the bandwagon*), locations (*build castles in the air*), recipients (*feed X to the lions*), sources (*start from scratch*), paths (*pay through the nose*), comitatives (*play with the big boys*), and instrumentals (*can cut X with a knife*). Experiencer and agent thematic roles, however, tend to be open slots, as do benefactives (*do a favor for X*).

This leads to the second observation. This is that fixed objects and PPs are common, but fixed subjects are rare. Instead, subjects are typically open. Crucially, there are no idioms or collocations with a fixed subject but an open object. Marantz (1984) noted this for idioms, and we can now see that it is true of collocations, as well. There are a few that include a subject (Pattern 9, especially), but none of them have an open slot for a selected argument. I will return in section 3.3.1 to the few potential counterexamples that exist (which all seem to be unaccusative verbs).

Similarly, and this is the third observation, in ditransitives, V X NP is well-attested (*give X a hand*), but V NP X is systematically missing, from both idioms and collocations (**give the little woman X*). This will be very important in sections 3 and 4.

Finally, I have identified collocations here as frequently co-occurring but literally interpreted combinations of words, whereas idioms involve combinations that are not interpreted literally. This is a simplification, because there are many cases where some of the words are interpreted literally, but others are not. Consider the following three examples that I classified above as idioms:

- (79) a. miss the boat ('missed an opportunity')
- b. foot the bill ('end up paying for something')
- c. Y steal X's thunder ('use someone else's ideas or words and take away their recognition')

In *miss the boat*, the object is not a literal boat, but the verb has its literal meaning. Nevertheless, the meaning of missing an opportunity only arises with this particular combinations of words (Riehemann 2001; see also Nicolas 1995, Schenk 1995). One cannot, for example, say that someone **caught the boat* to mean that they seized an opportunity. In *foot the bill*, the verb is not literal and only seems to be used as a verb in this one phrase. But *the bill* is completely literal. Nevertheless one cannot **foot the charges* or **foot the check*. Similarly for *Y steal X's thunder*, where the verb is used in its literal sense. One cannot instead say **Y rob X of X's thunder* or **Y swipe X's thunder*. The ability to substitute lexical items does not correlate with literal interpretations.

Going the other way, we also find some substitution with phrases that are not literal at all:

- (80) a. grasp/clutch/seize at straws
- b. give/get the boot
- c. beat/kick/knock the crap/shit/beejzus/tar/(living) daylights out of X

Some idioms allow multiple choices of certain verbs within a certain semantic range. Many conventionalized expressions also permit both *give* and *get*, as has been discussed heavily in the literature (e.g., Richards 2001, Harley 2002, Bruening 2010, Larson 2017).

I take this to show that there really is no sharp distinction between non-literal idioms and literal collocations. There are in-between cases, where part of the expression is literal and part is not. Importantly, there is nothing that correlates with literal versus non-literal interpretation. As we have seen, literal and non-literal

expressions overwhelmingly pattern together. I posit no theoretical distinction between literal collocations and non-literal idioms, and will call all such expressions *conventionalized expressions*, treating them as a single grammatical class.

3 Syntactic Constraints on Conventionalized Expressions

I turn now to syntactic constraints that have been proposed to hold of conventionalized expressions (idioms in particular), and show that many are incorrect. In contrast, the selection theory of Bruening (2010), with a minor modification, extends from idioms to collocations and accounts for many of the systematic patterns that we see.

3.1 Proposed Syntactic Constraint: Locality

I will begin with some recent proposals that there is a locality constraint on idiomatic interpretations. As I will show, there is no such constraint, and non-literal idioms can be quite large, even including multiple clauses. Collocations, on the other hand, tend to be smaller. This is exactly the opposite state of affairs from what these recent proposals would expect.

Two proposals regarding locality come from Marantz (1997) and Svenonius (2005). Marantz (1997) proposes that idiomatic interpretations are bounded by functional heads like *v*, the head that introduces the external argument of a verb. According to Marantz, this has the effect that any idiom with a fixed subject must be non-agentive. Idioms like *the shit hit the fan* are claimed not to have agentive subjects (which seems to be correct for this particular idiom). Idioms are also bounded by this functional head, so no idiom can go beyond the head that introduces the external argument.

Svenonius (2005) proposes that phase boundaries are absolute barriers to idiomatic interpretations. That is to say, all parts of a phrasal idiom must be minimally dominated by the same phasal node. In the phase theory of Chomsky (2000), the two phasal nodes in the clause are *vP* and *CP*.

Both proposals are clearly incorrect. There are numerous idioms that include an embedded clause with a null subject, and many of these null subjects are agentive:

- (81)
- a. NEG have a pot [PRO to piss in] ('lack any standing')
 - b. have an axe [PRO to grind] ('have a (quarrelsome) reason')
 - c. NEG have a leg [to stand on] ('have no standing')
 - d. a bitter pill [PRO to swallow] ('is unpleasant but must be accepted')
 - e. a tough row [PRO to hoe] ('is very difficult')
 - f. play hard [PRO to get] ('be coy')
 - g. kill the goose [that OP lays the golden eggs] ('destroy a profitable thing out of greed')
 - h. the straw [that OP broke the camel's back] ('seemingly minor action with major consequences')
 - i. rob Peter [PRO to pay Paul] ('discharge one debt by incurring another')
 - j. X cut off X's nose [PRO to spite X's face] ('do something self-destructive out of revenge')
 - k. teach X's grandmother [to suck eggs] ('offer advice to someone with more experience')

I indicate the null subject of a non-finite clause as *PRO*, and the null subject of a relative clause as *OP*. All of the above non-finite clauses seem to be agentive. They can all easily form passives and imperatives, for instance:

- (82) a. This pot was pissed in. Piss in that pot!

- b. The axe was ground. Grind that axe!
- c. The pill was swallowed. Swallow that pill!
- d. The garden has been hoed. Hoe those weeds!

All of the above idioms are also problematic for Svenonius's claim that idioms are bounded by phase boundaries, since they cross a CP boundary. Idioms can also include a full finite CP within them:

- (83)
- a. kill the goose [that lays the golden eggs] ('destroy a profitable thing out of greed')
 - b. X bite the hand [that feeds X] ('turn against someone you depend on')
 - c. the straw [that broke the camel's back] ('a seemingly minor thing that causes a disaster')
 - d. strike [while the iron is hot] ('act before it is too late')
 - e. close the stable door [after the horse has bolted] ('act after it's too late')
 - f. (Neg) count X's chickens [before they hatch] ('plan on something that might not happen')
 - g. X cross that bridge [when X comes to it] ('wait to act until it is necessary')
 - h. X bite off more [than X can chew] ('take on more than one can handle')
 - i. kick X [when/while X is down] ('treat someone badly who has already suffered a setback')
 - j. be that [as it may] ('that is true, but it does not change things')
 - k. [when hell freezes over] ('never')
 - l. [until the cows come home] ('for a very long time')

Since CPs are phase boundaries, these examples show that it is not necessary that all parts of an idiomatic phrase be contained within a single phase. (The last two include the CP phase and the vP phase.)

There are also idioms that obligatorily include high functional material in addition to a main verb, which would have to cross the vP phase boundary:

- (84)
- a. Break a leg! (obligatorily imperative; = 'give a good performance')
 - b. No can do. (Neg Modal V; = 'that's impossible')
 - c. can cut X with a knife (modal obligatory; = 'the feeling of X (e.g., tension) was strong and palpable')

One common instance of this is for negation to be included in the idiom:

- (85)
- a. NEG have a leg [to stand on] ('have no standing')
 - b. NEG have a pot [to piss in] ('have no standing')
 - c. NEG stop at anything ('be relentless')
 - d. NEG breathe a word about X ('keep X secret')
 - e. (can/could) NEG help it/Xself ('unable to stop oneself')
 - f. NEG see the forest for the trees ('focus on details and fail to see the big picture')

I will return to negation in section 3.2.3. Importantly, negation is not included in the lowest vP phase, so idioms that obligatorily include negation necessarily cross a phase boundary.

Harwood (2015) argues that certain idioms include progressive aspect. He gives the following examples (his example 83):

- (86)
- a. Something is eating X. ('something is obsessively bothering X')
 - b. X is pushing up daisies ('is dead')

- c. X is cruisin' for a bruisin' ('is acting in such a way as to invite retribution')
- d. X is chomping at the bit ('is impatient to start')

I find numerous examples on the internet of *push up daisies* and *chomp at the bit* in the simple past or in other forms besides the progressive. The other two do seem to only occur in the progressive (but *cruisin' for a bruisin'* might just to maintain the rhyme). In any event, Harwood argues that the progressive is included in the vP phase, and so these are not problematic for the claim that idiomatic interpretation is limited to a single phase. However, he also cites an idiom that seems to include perfect aspect, in (87a). To this we can add the two in (87b–c), from Lebeaux (2009: 65):

- (87)
- a. X has been around the block a few times ('X is very experienced')
 - b. X has had it ('X is done/has lost all patience')
 - c. X has gone to the dogs ('X has declined in quality')

All of these obligatorily include the perfect aspect, which would have to be outside the vP phase on Harwood's assumptions. There does not seem to be a difference between progressive (one to two idioms) and perfect (three idioms) in this regard; in fact there are *more* idioms that include the perfect than include the progressive.

From the above data, we can see that idioms can cross the vP phase boundary by involving modals, aspect, negation, imperatives, and other functional material in addition to lexical material inside the VP. They can also embed full CPs, meaning that they can cross the CP phase boundary. They can also include agentive subjects (typically null ones of embedded clauses). The specific claims of Marantz (1997), Svenonius (2005), and Harwood (2015) regarding a strict locality constraint on idioms are not borne out by the data.

It is also not clear why there would be a locality constraint on special interpretations. Numerous other aspects of interpretation must be able to involve agents and must be able to cross the vP phase boundary. For instance, an agent must be interpreted with respect to a lexical verb. Aspect, mood, negation, and so on must be able to operate on VP material. Association with focus can certainly cross phase and clause boundaries. So can quantifier scope. Control is a syntactic and semantic relation that relates a higher NP with the subject of an embedded clause. Why should idiomatic interpretations be different, such that only they cannot cross phase boundaries or the head v? (Note that the longer idioms shown in this subsection are just combinations of the patterns identified in section 2.2: *NEG have a pot to piss in* is the combination of a functional element plus a main verb (pattern 7) plus a verb and its object (pattern 2a), plus a modifier of the N (like pattern 1f but with an infinitive, like pattern 5); the modifier of N is a complete clause (pattern 9).)

Conventionalized expressions tend to be small, but this is only a tendency. In fact, there seems to be a real difference between non-literal idioms and purely literal collocations in this regard. I have found very few collocations that consist of more than two or three constituents. In contrast, this section has illustrated quite a few idioms that include multiple clauses. This is exactly the opposite of the state of affairs from what would be expected on the view that idiomatic, non-literal interpretations are the ones that are bounded by a locality constraint.

I suspect that the difference between collocations and idioms in this regard is due to how they develop. Collocations are simply combinations of two or three words or phrases that get repeated until they become frequent and conventionalized. Many idioms, in contrast, derive from sayings, proverbs, and quotations from literature. As such, they start out quite long, often as entire sentences, and get shorter as they become conventionalized (see Moon 1998: 115 and section 3.3 below). Given this difference in etiology, it is even more striking, and more in need of a syntactic explanation, that the syntactic forms of idioms and collocations are so similar. I explore what the syntactic constraints that operate on them are in the next subsection.

3.2 Constraints on Structure and the Selection Theory

We have now seen that there is no locality constraint on conventionalized expressions. However, we have also seen that conventionalized expressions take only certain forms and not others. In the rest of this section, I show that the selection theory of Bruening (2010) can explain the attested patterns of collocations and idioms alike.

3.2.1 Previous Proposals Regarding Idioms

As noted above, it has been common over the years for researchers to propose that idioms are underlying constituents, excluding all non-idiomatic material (e.g., Larson 2017). As also noted above, this view was shown long ago to be untenable (e.g., Ernst 1981, Nunberg et al. 1994, Nicolas 1995, O’Grady 1998). Idioms can be disrupted by modifiers (88); the determiner that usually appears can be exchanged for a different one (89, where the canonical form is *beat around **the** bush*); numerous idioms have an open slot for a possessor (90); and many idioms consist of a verb, NP, and P, but exclude the object of that P (91):

- (88) Before the symbolic shit hit the colonial fan recently at the University of Cape Town, . . .
(<http://www.thejournalist.org.za/spotlight/we-love-uct-says-student-who-covered-rhodes-in-shit>)
- (89) I’ve seen you beat around more bushes than a landscaper.
(<http://www.kungfumagazine.com/forum/showthread.php?42550-OT-9-11-Truth-Symposium/page5>)
- (90) lose X’s cool, get X’s goat, fill X’s shoes, . . . (O’Grady 1998: (4c–e))
- (91) beat the bushes for X, cast a pall on X, NEG hold a candle to X

In none of the cases above does the idiomatic material constitute a constituent that excludes all non-idiomatic material.

Researchers have made various other proposals, as well. Baltin (1989) proposes that idioms involve only the head of a phrase and the head of one of its complements. This is not correct, because specifiers can be involved too, as can modifiers (e.g., *play devil’s advocate*, *beat/flog a dead horse*). van Gestel (1995) proposes that idioms consist of the selection of a lexical head by another lexical head. This is also not correct, for the same reason: idioms can include functional material, like negation, aspect, and imperative mood, as was shown in section 3.1. Koopman & Sportiche (1991: 224) propose a constraint that says that if X is the minimal constituent containing all the idiomatic material, the head of X is part of the idiom. This is correct, but it is too weak. For instance, it would permit any idiom that consisted of just the head C of a CP plus any subset of material within the CP, including discontinuous material. There could be idioms like, *if the devil Vs X* or *if X Vs the wolves*, with open material except for the subject in the first case and the object in the second. No such idioms exist. Finally, Harley & Jung (2015: 724) suggest that the head of an idiomatic constituent must combine with its sister as an idiom before idiomatizing with other constituents. This is also clearly false, as we have seen numerous examples where a preposition combines with other constituents as an idiom, leaving its own sister out (91). There are also idioms that include a verb and an adjunct PP, but which exclude the object of the verb (*nip X in the bud*); if arguments always combine with verbs before adjuncts do, then these also counterexemplify Harley and Jung’s proposed constraint.

3.2.2 The Selection Theory

Bruening (2010), building on O’Grady (1998), proposes the following as an account of the syntactic forms that idioms can take:

- (92) The Principle of Idiomatic Interpretation:
X and Y may be interpreted idiomatically only if X selects Y. (Bruening 2010: 532, (24))

In this formulation, idiomatic interpretation is dependent on selection. Idioms consist of one element selecting another, and the two may then be interpreted as an idiom. Note that here selection is for particular lexical items (the “I-selection” of Pesetsky 1992).

To illustrate how this works, one of the most common type of phrasal idiom consists of a verb and its direct object, for instance *buy the farm* (‘die’). This idiom consists of a verb, *buy*, which selects an NP object headed by *farm*. This follows the Principle of Idiomatic Interpretation. Another common idiom pattern consists of a verb and its object, but with an open slot for the possessor of the object. Possessors can be open slots because they are not selected. In *get X’s goat*, for instance, the verb *get* selects an NP headed by *goat*, and so the two can be interpreted idiomatically. The possessor is not selected and does not disrupt this selection. (If we wished to capture the fact that in most uses of this idiom, a possessor is included, we can do that by saying the head N selects a possessor, but does not specify what it is.)

Modifiers can also be added to idioms, as in *kick the filthy habit*, because they do not disrupt selection. In this example, *kick* selects *habit*, and a modifier can be added without disrupting this selection. Similarly, in *beat around the bush* being used as *beat around more bushes than a landscaper*, the determiner can change because it is not selected. Determiners are never part of idioms, as Bruening et al. (2017) show, and can change depending on the context. (Note that this requires the rejection of the DP Hypothesis, which says that the head of an NP is not N but D. This hypothesis is incompatible with the facts of idioms. See Bruening et al. 2017. We also have to reject the view that a Deg head is the head of a comparative construction; in this example, the head of the object *more bushes than a landscaper* must be the N *bush*.)

Note that modifiers *can* be part of an idiom, as in *beat/flog a dead horse*. This is possible because modifiers select what they modify (Pollard & Sag 1994). This is made clear by the difference between adverbs and adjectives, where adjectives adjoin only to nominal projections while adverbs adjoin to other categories. This could only be an instance of categorial selection, but it would not be correct to say that the adjoined-to category selects the adjunct. It must therefore be that adjuncts select the categories they adjoin to (and projection is not determined by selection; see Bruening 2010: 533–534). In *beat/flog a dead horse*, the verb selects an NP headed by *horse*, and the adjective *dead* does so as well. This is consistent with the Principle of Idiomatic Interpretation, and so the whole can have an idiomatic interpretation.

Some idioms permit more than one word, like *beat or flog a dead horse*; *clutch at straws* is also attested as *grasp at straws* and *seize at straws*. L-selection can specify a list of lexical items in addition to a single lexical item, or even a semantic class of lexical items, as in “families of idioms” like *beat the crap out of*. Here the verb can be one of several verbs like *beat* and *kick*, and the noun can also vary within a limited range, as *crap*, *shit*, *tar*, etc. A particular verb or class of verbs can select a list of particular lexical items, or a class of lexical items, with the result that the particular words chosen can even be somewhat productive (see McGlone et al. 1994). (It is likely that in some cases, like *clutch/grasp/seize at straws*, speakers store an abstract semantic concept rather than a particular verb, and they will then use any of the small set of verbs that match that concept.)

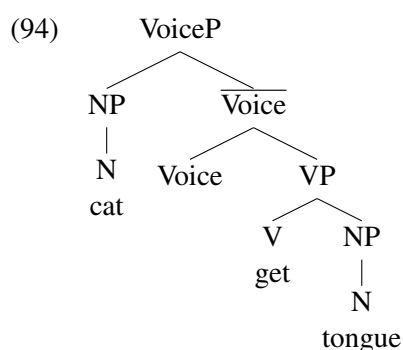
There are other proposals that are similar to the selection theory. O’Grady (1998) formulates a very similar approach in terms of Dependency Grammar. In O’Grady’s formulation, this is mostly equivalent to the selection theory. However, it has been altered slightly in the notion of the *catena* in Osborne et al. (2012). I will not adopt this theory, because it introduces a syntactic notion, the *catena*, that I believe syntactic theory can do without. Similarly, another approach posits phrasal *constructions* as primitives in grammar and views phrasal idioms as examples of these constructions (e.g., Riehemann 2001). I also believe that syntactic theory can and therefore should do without the formal notion of a construction. In contrast, every theory needs selection in some form or other.

3.2.3 Building Idioms through Selection

As was described briefly above, phrasal idioms can include more than two items by building a chain of selection. So long as each item either selects or is selected, the whole can constitute an idiom. I will go through a few of the longer idioms from above to illustrate the workings of the system, specifically the following:

- (93) a. The cat got X's tongue. ('X is speechless')
 b. NEG have a pot [PRO to piss in] ('lack any standing')
 c. X cross that bridge [when X comes to it] ('wait to act until it is necessary')

The idiom *the cat got X's tongue* seems to involve a verb, a subject, and an object. I adopt the hypothesis from Kratzer (1996) that external arguments are selected by a head Voice and are projected in that head's specifier:



Objects are selected by the lexical verb and merged as its complement, as shown.

I will assume that non-selected material is generally irrelevant to the idiom. Many functional elements are also not part of idioms, for instance the definite article. Possessors may or may not be; here I assume it is not (note that many instances of this idiom without the possessor, for instance *cat got the tongue*, are attested). In the tree above, Voice selects an NP headed by *cat* and a VP headed by *get*. *Get* selects an NP headed by *tongue*. The entire phrase satisfies the Principle of Idiomatic Interpretation, and so it can be interpreted idiomatically. (Again, if we wished to capture the fact that in most uses of this idiom, a possessor is included, we can do that by saying the head N selects a possessor, but does not specify what it is. Alternatively, an unspecified possessor as an optional modifier selects the N.)

In the idiom *NEG have a pot to piss in*, the non-finite clause is probably an infinitival relative modifying *pot*. As adjuncts, relative clauses select phrases of category N to adjoin to. Here, the relative clause selects a nominal headed by *pot*. Within the infinitival relative clause, I assume that a null C selects infinitival *to*, which in turn selects a VoiceP. Voice selects PRO as its specifier and a VP headed by *piss*. *Piss* selects a PP headed by *in*, while *in* selects a null operator which moves to Spec-CP. In the main clause, the verb *have* selects an NP headed by *pot*. As for negation, I will assume that it is a feature that can be specified on various syntactic categories, including clauses and VPs. This [Neg] feature is spelled out in different ways; in English, it is often spelled out as the adverb *not* adjoined to the highest VP of the clause. Idioms with negation often surface in different forms; for instance, *leave no stone unturned* is also robustly attested as *NEG leave any stone unturned*; the same is true for various other idioms with negation and an existential (e.g., *stop at nothing* ~ *NEG stop at anything*). Negative quantifiers are the combination of an existential and this abstract NEG (see, e.g., Penka 2012).³

³Another option is to view obligatorily negative idioms as negative polarity items. Then NEG would not necessarily be part of the idiom, but the idiom would be specified as having to occur in the scope of negation. This is already necessary for certain NP

Finally, we have the interesting example of *X cross that bridge when X comes to it*, where the subject of the embedded clause is obligatorily covalued with the matrix subject. I assume that the idiom includes matrix Voice, which selects an external argument but does not specify what it is. The embedded clause adjoins to a projection of Voice and selects Voice, further specifying that the selected external argument of Voice must control the subject of the embedded clause. I will assume that this is an example of a switch-reference system, which must be mediated by C (e.g., Finer 1985). C selects both the higher Voice, and through a chain of selection (C-T-Voice) selects the subject embedded within it. Note that the embedded object pronoun *it* is also obligatorily covalued with a matrix argument (the object); this will also have to be specified.

3.2.4 Impossible Idioms

So far we have seen how the selection theory accounts for existing idioms. It also rules out certain types of phrases as idioms. In particular, it rules out discontinuities in idiomatic phrases, where X selects Y and Y selects Z but only X and Z are part of the idiom. We saw the example of *X has gone to the dogs* above, where the idiom consists of perfect aspect selecting a verb which selects a PP headed by *to* which selects an NP headed by *dogs*. An impossible idiom would have perfect aspect included plus something selected by the main verb, but the main verb is not part of the idiom and is free to vary. This would be something like, **X has V-en an old shoe*, perhaps with a meaning like, ‘X has V-en every conceivable thing’. There is no such idiom, and the selection theory accounts for this.

Similarly, there could be no idiom that includes negation, imperative mood, and an object, but where the verb is free to vary. This would have a form like **NEG V! the golden goose*, with a meaning like, ‘don’t V a profitable thing’. Again, no such idiom exists. The selection theory also rules out an idiom where the verb is fixed as is an adjective modifying the object, but the object itself can vary: **beat a dead X*, **be in hot X*. The verb does not select the adjective and the adjective does not select the verb, so such idioms violate the Principle of Idiomatic Interpretation. Also ruled out is an idiom consisting of a verb and the possessor of the object, but not the object itself: **play the devil’s X* (O’Grady 1998: 287). Again, the verb does not select the possessor and the possessor does not select the verb. (In *play devil’s advocate*, the verb selects the object, and the possessor is selected by the head noun *advocate*, as it is a selected argument of that noun.)

Another thing to note is that, in all the examples of relatively long idioms in section 3.1 with an embedded clause, the constituents of the embedded clause are almost entirely fixed. They can only co-vary with constituents in the matrix clause (for instance, the subject in *X cross that bridge when X come to it*, and also tense: *we will cross that bridge when we come to it* versus *we crossed that bridge when we came to it*). This is as predicted by the Principle of Idiomatic Interpretation: building a long idiom requires a long chain of selection, with the result that all selected elements have to be included as part of the idiom. Only non-selected elements could ever be added or changed (the exact values can also be dependent, as for instance tense, but tense itself would be selected by C in the embedded clause).

3.3 The Constraint on Idiomatic Interpretation

In addition to the Principle of Idiomatic Interpretation, Bruening (2010: 532, (25–26)) proposes a constraint on possible idioms, as follows:

idioms, like *a living soul*, which can occur with different verbs (*don’t tell a living soul*, *I didn’t see a living soul*) but must occur with negation. Negative verbs also license it: *I doubt that there’s a living soul out there...*, so sentential negation is not a crucial component of this idiom. Andrew Murphy points out that obligatorily negative idioms can occur with NEG raising, as in *I don’t think he has a pot to piss in*. If NEG is indeed a selecting part of the idiom, then this would necessitate a movement derivation of NEG raising, as in Collins & Postal (2014). If negative idioms are instead negative polarity items, then it would not.

- (95) Constraint on Idiomatic Interpretation:
If X selects a lexical category Y and X and Y are interpreted idiomatically, all of the selected arguments of Y must be interpreted as part of the idiom that includes X and Y.
- (96) Lexical categories are V, N, A, Adv.

This constraint rules out certain patterns: having a fixed subject and verb but an open slot for an object; and the missing ditransitive pattern, *V NP X.

3.3.1 Fixed Subject and Verb with Open Object

Marantz (1984) noticed that, while there are many idioms with a fixed verb and object and an open slot for the subject (e.g., *X buy the farm*), there are no idioms with a fixed verb and subject and an open slot for an object (**the cat got X*). This follows from the Constraint on Idiomatic Interpretation in (95), plus the VoiceP analysis of Kratzer (1996). If the external argument is selected not by the verb but by Voice, then such an idiom would have to include Voice and its selected specifier. Voice would also select the VP, in order to include the main verb, but then all of V's arguments would have to be included, as well, given the Constraint on Idiomatic Interpretation. This permits idioms where all three are idiomatic, like *the shit hit the fan* (or embedded clauses as in *kill the goose [that OP lays the golden eggs]*), but rules out the missing pattern where the object is open.

Various examples that have been advanced as potential counterexamples turn out not to be real counterexamples. For instance, *a little bird told X Y* is not actually an idiom with a fixed subject and verb. The meaning resides entirely in the NP *a little bird*, meaning 'an anonymous source of information'. This NP can occur with other verbs of communication, as in *a little bird is broadcasting that, a little bird whispered that to me, I heard it from a little bird*, or even with no verb of communication, as in, *Had Varys's little birds failed him for once?* (George R.R. Martin, *A Clash of Kings*). Similarly, *the lovebug bit X* is not a phrasal idiom with a fixed subject and verb, as the NP *the lovebug* can occur in many different expressions: *the lovebug has struck Bill, Bill has the lovebug* (and *lovebug* can also vary, with other things replacing *love*). *The devil drove X to it/Y* is also not a phrasal idiom with a fixed verb and subject and an open object, as it also occurs in various different forms: *the devil made him do it, the devil forced him to do it, the devil got into him*.

There are a few potentially real counterexamples, below:

- (97) something/what is eating X, the bottom fell out of X, the roof caved in on X

However, these verbs are all unaccusative. This is clearly true of *fall out of* and *cave in on*. It is not obviously true of *eat*, but in the meaning of this idiom, the verb is interpreted as a psych verb, meaning something like 'bother (obsessively or to an extreme)'. Psych verbs are known to behave unusually, and have been analyzed as having no subject underlyingly. Instead they have two objects, one of which becomes the surface subject (e.g., Belletti & Rizzi 1988). If this is correct, then in all three of these idioms, both arguments are arguments of the main verb, and neither is an argument of Voice. The idiom is simply the V and one of its two selected arguments. Since nothing that selects V is part of the idiom, then not all of V's arguments need to be part of the idiom.

A common development of proverbs into idioms is also accounted for by the Constraint on Idiomatic Interpretation. Many proverbs are complete clauses. Moon (1998: 115) notes that proverbs that are full clauses often lose their subject to become VP idioms (note that determiners can also change, from *a straw* to *straws*):

- (98) a. A drowning man will clutch/grasp at a straw → to clutch/grasp at straws
b. Don't look a gift horse in the mouth → to look a gift horse in the mouth

- c. Don't put the cart before the horse → to put the cart before the horse
- d. You can't have your cake and eat it, too → to have X's cake and eat it, too
- e. Make hay while the sun shines → to make hay
- f. The pot can't call the kettle black → to call the kettle black

It never happens that the object becomes the open slot, while the subject is retained. In my judgment, turning *the pot can't call the kettle black* into *call the kettle black* is perfectly natural, but turning it into *the pot can't call X black* is entirely unnatural. Hence the Constraint on Idiomatic Interpretation captures something important about native speaker competence.

Note that this asymmetry could not be about subject-predicate relations, where the subject of a predicate prefers to be the open slot. Underlying objects can be turned into subjects of predicates by using the passive, but *X is called black by the pot* is just as unnatural an idiom derived from the proverb. This follows from the Constraint on Idiomatic Interpretation combined with a transformational view of the passive: If the idiom only occurs in the passive, then Passive Voice must be part of the idiom (see Kratzer 1996, Bruening 2013). It selects VP, but then all the selected arguments of the V must be included, ruling out the underlying object/surface subject being an open slot.

3.3.2 Functional Categories: Object of P

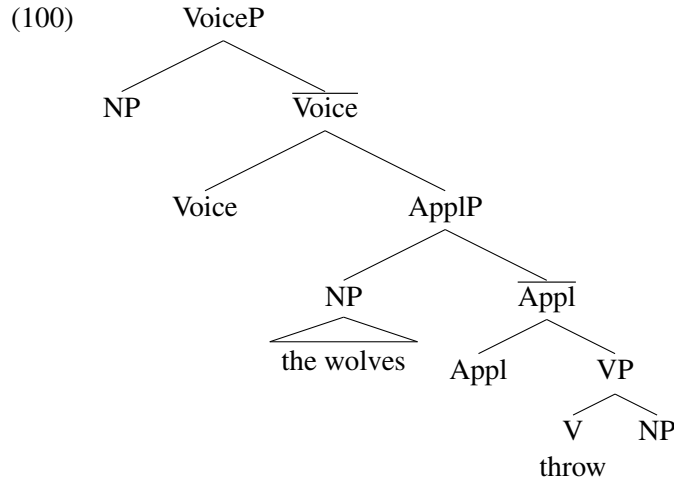
The Constraint on Idiomatic Interpretation in (95) refers to *lexical* categories. The reason to exclude functional categories is that, as O'Grady (1998: 300–301) notes, there are numerous idioms that consist of a verb plus a PP, but where the object of the P is an open slot (many of them also include a direct object of the verb):

- (99) beat the bushes for X, carry a torch for X, cast a pall on X, do a number on X, give it to X, give rise to X, give the lie to X, give voice to X, have a bone to pick with X, have an axe to grind with X, NEG hold a candle to X, light a fire under X, make head or tail of X, pay lip service to X, pull the plug on X, shed/cast light on X, sink Y's teeth into X, take a back seat to X, take a gander at X, throw the book at X, turn a blind eye to X

In these idioms, the verb selects an NP and a PP headed by a particular preposition. The complement of P does not have to be included, given the Constraint on Idiomatic Interpretation, because P is not a lexical category.

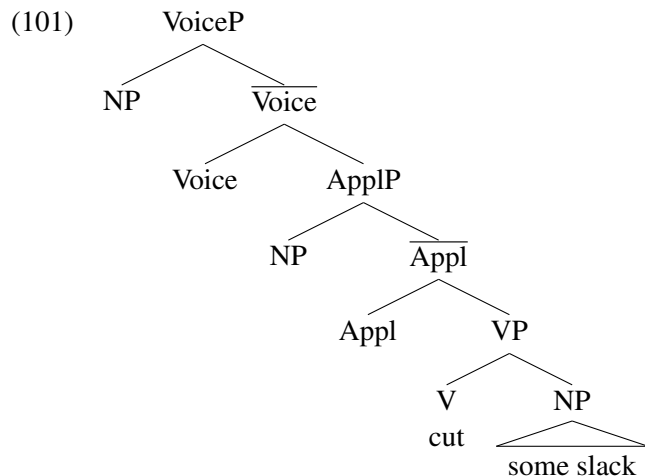
3.3.3 Missing Double Object Pattern

The Constraint on Idiomatic Interpretation in (95) also rules out the missing double object pattern, *V NP X, given a particular analysis of double object constructions (more on this in section 4). Consider a potential but non-occurring idiom like **throw the wolves X* (cf. *throw X to the wolves*). Bruening (2001, 2010) proposes the following structure for double object constructions, where the first object is projected in the specifier of an Appl(icative) Phrase between Voice and VP (modifying a proposal in Marantz 1993):



In order for an idiom to include the NP in Spec-AppIP, it would also have to include Appl. Appl would select the NP in its specifier, and its complement VP. This satisfies the Principle of Idiomatic Interpretation. However, by the Constraint on Idiomatic Interpretation, all the selected arguments of V would have to be included, as well (as in *give the devil his due*). This is why there is no idiom with a fixed verb and first object, but an open slot for the second object.

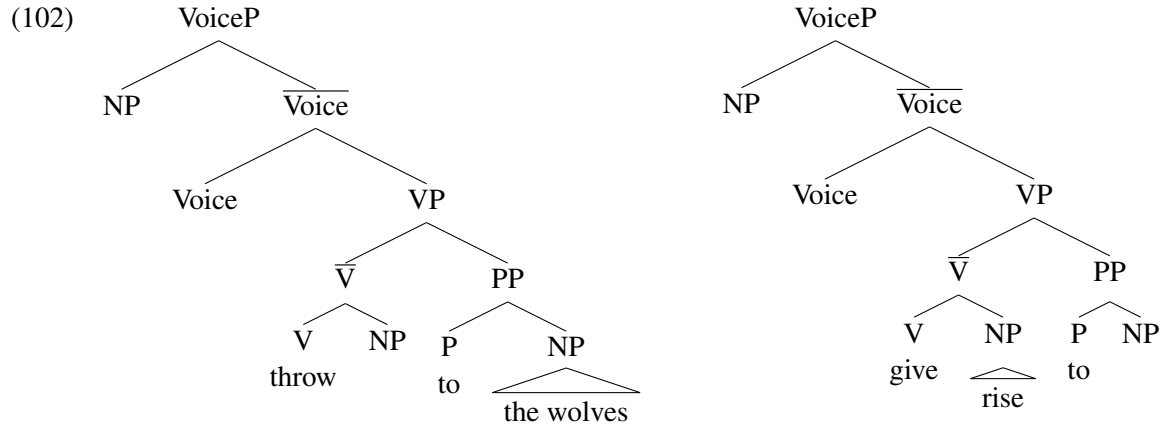
In contrast, the common double object pattern V X NP satisfies the Constraint on Idiomatic Interpretation:



Here the verb *cut* selects an NP headed by *slack*, and by the Principle of Idiomatic Interpretation, they can form an idiom together. The open first object is projected in the specifier of AppIP and is not included. The Constraint on Idiomatic Interpretation is not violated. (If the idiom only occurs in the double object construction and not also the prepositional dative construction, as *cut X some slack* seems to, then we might want to say that the head Appl is also part of the idiom. Its specifier is not. This also satisfies the Constraint on Idiomatic Interpretation.)

I will adopt the following structure for the prepositional dative construction, where the verb selects both the direct object and the PP:⁴

⁴This structure is most consistent with constituency tests: some speakers permit the verb and direct object to act as a constituent, excluding the PP (see Janke & Neeleman 2012). It is also consistent with the binding facts discussed in Barss & Lasnik (1986), given that binding depends not on c-command but on precede-and-command (Bruening 2014). Note also that this is the structure I assume for all V NP PP frames, regardless of the semantics of the PP. See section 4, especially section 4.5.



In this structure, it is possible to have idioms of both the form V X PP (*throw X to the wolves*) and V NP P X (*give rise to X*), because the verb selects both the NP and the PP. Since nothing that selects V is part of the idiom, not all of the selected arguments of the verb need to be included in the idiom. Since the P is a functional element, its object also does not need to be included. It is also possible to have idioms consisting of the verb and both of its arguments, as in *give hostages to fortune*.

As can be seen, the ApplP analysis of ditransitives in Bruening (2001, 2010) combined with the Constraint on Idiomatic Interpretation in (95) succeeds in ruling out the missing *V NP X pattern and in accounting for the patterns of idioms in ditransitives generally. As we will see in section 4, it is the only analysis that does.

3.4 Generalizing to All Conventionalized Expressions

Having outlined the selection theory of idioms, we can ask whether this theory can be extended to literal collocations as well, given that the empirical study reported above found that they behave alike. It does, quite simply. We can generalize the Principle of Idiomatic Interpretation to cover all conventionalized expressions as follows:

- (103) The Principle of Conventionalization:
X and Y may become a conventionalized expression only if X selects Y.

We could also generalize the Constraint on Idiomatic Interpretation to the following (to be revised):

- (104) Constraint on Conventionalized Expressions (to be revised):
If X selects a lexical category Y and X and Y are a conventionalized expression, all of the selected arguments of Y must be part of the conventionalized expression that includes X and Y.
- (105) Lexical categories are V, N, A, Adv.

Let us see how this fares with the collocation patterns that were identified in section 2, beginning with modification in Pattern 1:

- (106) *Pattern 1: Modification*
- Adj N: honest mistake (idiom: cold feet)—Adj selects N
 - Adv Adj: acutely aware (idiom: writ large)—Adv selects Adj
 - Adv V: go smoothly (idiom: only have eyes for X)—Adv selects V
 - Adv Adv/P: hardly ever, way ahead (no idioms)—Adv selects Adv/P
 - N PP: call of duty (idiom: elephant in the room)—PP selects N

- f. Adj PP/NP: missing in action, worth a fortune (idiom: running on fumes)—Adj selects N or PP selects Adj
- g. Comparison: better than a kick in the pants (idiom: X's bark is worse than X's bite)—Deg selects Adj and NP

As can be seen, these all involve a local selectional relation, exactly as the Principle of Conventionalization (103) requires. The same is true of Pattern 2, involving a verb and its arguments/adjuncts:

(107) Pattern 2: Verbs and Arguments/Adjuncts

- a. V NP: answer the door (idiom: back the wrong horse)—V selects NP
- b. V X's NP: break X's heart (idiom: clip X's wings)—V selects NP
- c. V NP P X: break the news to X (idiom: beat the bushes for X)—V selects NP and P
- d. V X NP: give X permission, do X a favor (idiom: cut X some slack)—V selects NP
- e. V NP NP: give it a rest (idiom: give the devil his due)—Appl selects NP and VP, V selects NP
- f. V NP PP: Y bring a tear to X's eye (idiom: build castles in the air)—V selects NP and PP (or adjunct PP selects VP)
- g. V PP: come into view (idiom: beat around the bush)—V selects PP or PP selects V
- h. V X PP: bring X to justice (idiom: feed X to the lions)—V selects PP or PP selects V
- i. V PP P X: keep in touch with X (no idioms)—V selects PP and P
- j. V P... X: NEG bet on X (no idioms)—NEG selects V, V selects P
- k. V (NP) Adj: call in sick, make a long story short (idiom: getting busy, get X's feet wet)—V selects (NP and) Adj
- l. V X Adj: drive X crazy (idiom: pull it/X together)—V selects Adj or Adj selects V
- m. V (NP/X) Infinitive: lead X to believe (idiom: let sleeping dogs lie)—V selects (NP and) infinitive C, C selects T selects Voice selects NP and VP
- n. V (NP) CP: tell me it isn't/ain't so (idiom: X teach X's grandmother to suck eggs)—V selects (NP and) CP, C selects T selects Voice selects NP and VP

Patterns 3–10 also fall into line:

- (108) (P3) Coordinations: law and order (idiom: the old ball and chain)—Conj head selects NP and NP
- (109) (P4) P NP: against the law (idiom: behind the eight ball)—P selects NP
- (110) (P5) Adj Infinitive: hard to say (idiom: fit to be tied)—Adj selects infinitive C, C selects T selects Voice selects NP and VP
- (111) (P6) Predicative: no money down (idiom: back against the wall)—Predicate selects NP subject
- (112) (P7) F V (NP P) X: can't stand X (idiom: has had it)—NEG selects Mod, Mod selects V (idiom: Asp selects V, V selects NP)
- (113) (P8) V (X's) NP Y: X spend X's life Y-ing (idiom: X cut X's teeth Y-ing)—V selects NP and nonfinite clause/gerund
- (114) (P9) Full Clause: easy does it (idiom: the shit hit the fan)—Voice selects NP and VP, V selects NP
- (115) (P10) Covert Conditional: easy come, easy go (idiom: monkey see, monkey do)—covert *if* (C) selects main clause and T in its own clause, chain of selection in each clause

As can be seen, all of the collocation and idiom patterns identified here involve continuous chains of selection, exactly as the selection theory requires. The Principle of Conventionalization in (103) covers the data very well.

An important question that I will not attempt to answer here is *why* selection would be implicated in conventionalized expressions, literal and non-literal alike. I speculate that selection has something to do with how the human mind builds and stores complex expressions. However, I will have to leave for future research a full exploration of this issue. What is important here is that the syntactic relation implicated in the syntactic forms that conventionalized expressions take is selection.

3.4.1 Problems for the Constraint on Conventionalized Expressions

While the Principle of Conventionalization in (103) extends very well to literal collocations, there are a few problems for the Constraint on Conventionalized Expressions in (104). There are several conventionalized expressions that include a chain of selected elements, but then allow an open slot in a fairly low position. These include the following:

- (116)
- a. well worth X
 - b. barely able to X, recommend X highly, strongly support X, take X seriously
 - c. lead X to believe Y, let X know that Y
 - d. can't afford X, can't help X-ing, can't stand X
 - e. get the impression that X, the fact remains that X
 - f. enormous amount of X
 - g. can cut X with a knife, nip X in the bud

The problem with these expressions is that one element of the expression selects a lexical category, but then not all of that lexical category's selected arguments are included in the conventionalized expression. This violates the Constraint on Conventionalized Expressions in (104).

Consider for example the idiom *can cut X with a knife*. This always occurs with the modal *can*, and it has an adjunct PP. Suppose the modal selects VoiceP. The PP either selects VoiceP (Bruening 2013) or VP. Either way, VP is also selected, by Voice if by nothing else. Being a lexical category, all of its selected arguments have to be part of the conventionalized expression, but they are not.

The expressions listed in (116) that have an open clausal argument might not be problems. For instance, in *get the impression that X*, we could say that the N *impression* selects a CP headed by *that*. This is not a lexical category, so not all of the selected arguments of *that* need to be included. The same could be said of *let X know that Y*, while *enormous amount of X* has the open slot as the object of a preposition, which is functional. However, many of the open slots are objects of verbs, as in *can cut X with a knife*, *nip X in the bud*, *recommend X highly*, etc. These definitely violate the Constraint on Conventionalized Expressions in (104).

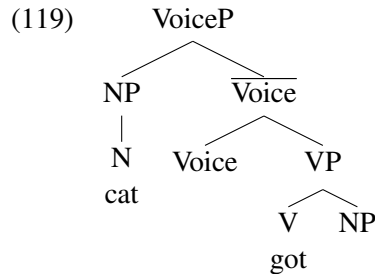
We obviously do not want to simply get rid of the Constraint on Conventionalized Expressions, since it is what accounts for the patterns that are striking in their absence: *NP V X (fixed subject, open object), and *V NP X (the missing ditransitive pattern).

3.4.2 Reformulation of the Constraint

Looking at the exceptions in (116), what they have in common is that there is no fixed NP that is hierarchically higher than the open NP slot. The two missing patterns that we want to rule out, *NP V X and *V NP X, both have a fixed NP that is hierarchically higher than the open NP slot. I therefore propose that we reformulate the Constraint on Conventionalized Expressions to the following:

- (117) Constraint on Conventionalized Expressions (revised):
 A conventionalized expression that includes an NP X as part of the conventionalized expression may not have an open NP argument slot Y if X is structurally superior to Y.
- (118) X is structurally superior to Y iff the set of nodes that dominate X is a proper subset of the set of nodes that dominate Y.

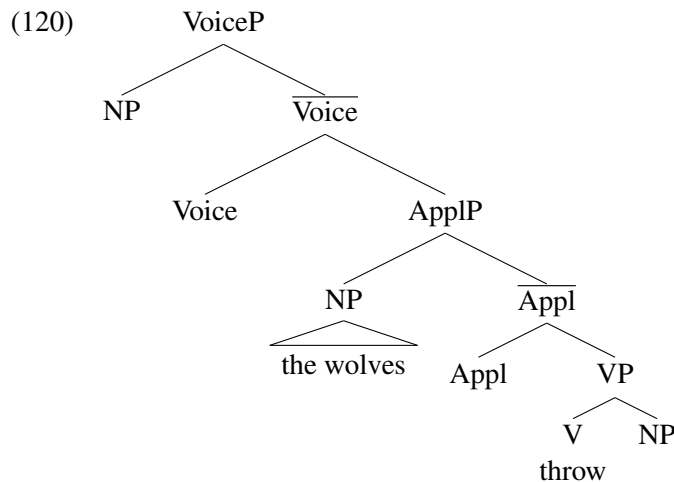
This still rules out a fixed-subject, open-object idiom like **cat got X*:



In such an expression, Voice selects the NP *cat*, and it also selects a VP headed by *get*. This permits all three to form a conventionalized expression, by the Principle of Conventionalization. However, the revised Constraint on Conventionalized Expressions in (117) rules this out: the set of nodes that dominate the NP *cat* ($\{\text{VoiceP}\}$) are a proper subset of the set of nodes that dominate the open object NP ($\{\text{VP}, \overline{\text{Voice}}, \text{VoiceP}\}$).

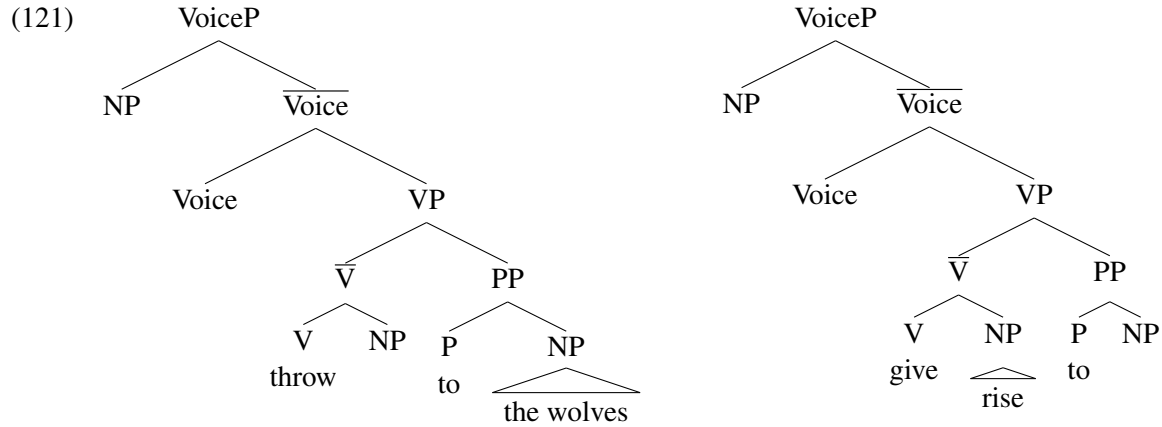
An open possessor is still allowed, as in *cat got X's tongue*, because the possessor is not an *argument*. As formulated, the Constraint on Conventionalized Expressions in (117) only refers to open *argument* slots (i.e., things that are selected, in keeping with the selection theory).

The missing double object pattern (**throw the wolves X*) is also still ruled out:



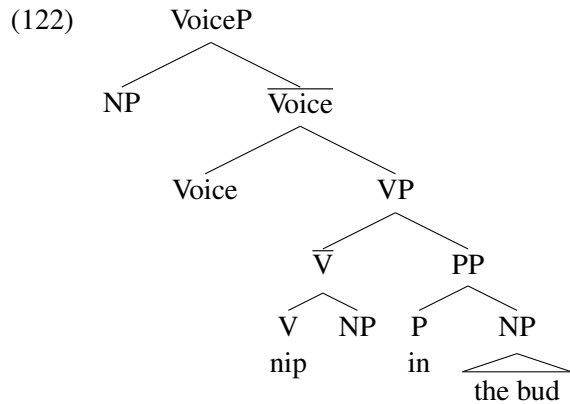
Here, the set of nodes that dominate the NP *wolves* ($\{\text{ApplP}, \overline{\text{Voice}}, \text{VoiceP}\}$) is a proper subset of the set of nodes that dominate the open NP object ($\{\text{VP}, \overline{\text{Appl}}, \text{ApplP}, \overline{\text{Voice}}, \text{VoiceP}\}$). If both NPs are included in the conventionalized expression, as in *give the devil his due*, the constraint is not violated.

The revised constraint also still allows both prepositional dative patterns:



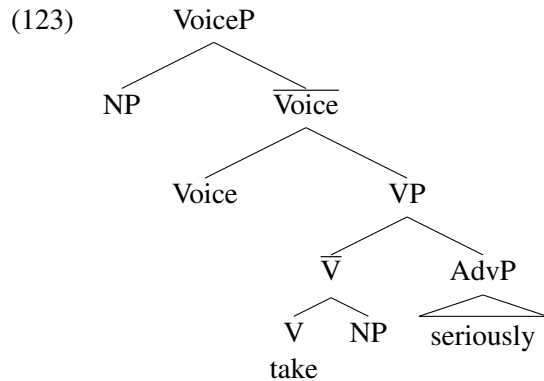
In the tree on the left, the set of nodes that dominate *the wolves* ($\{\text{PP, VP, } \overline{\text{Voice}}, \text{VoiceP}\}$) is not a proper subset of the set of nodes that dominate the open NP ($\{\overline{\text{V}}, \text{VP, } \overline{\text{Voice}}, \text{VoiceP}\}$). In the tree on the right, the set of nodes that dominate the NP *rise* ($\{\overline{\text{V}}, \text{VP, } \overline{\text{Voice}}, \text{VoiceP}\}$) is not a proper subset of the set of nodes that dominate the open NP ($\{\text{PP, VP, } \overline{\text{Voice}}, \text{VoiceP}\}$). (Of course, both NPs can be included without violating the constraint, as in *give hostages to fortune*.)

An adjunct PP is also permitted by the constraint:



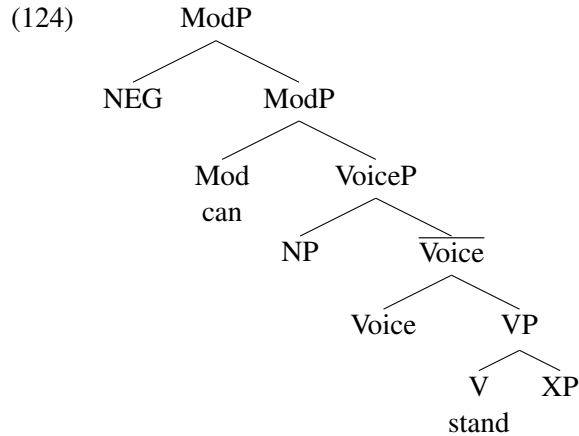
The set of nodes that dominate the NP *the bud* ($\{\text{PP, VP, } \overline{\text{Voice}}, \text{VoiceP}\}$) is not a proper subset of the set of nodes that dominate the open NP object ($\{\overline{\text{V}}, \text{VP, } \overline{\text{Voice}}, \text{VoiceP}\}$).

The revised constraint also permits a collocation that consists of a verb and an adverb but with an open NP object, like *take X seriously*:



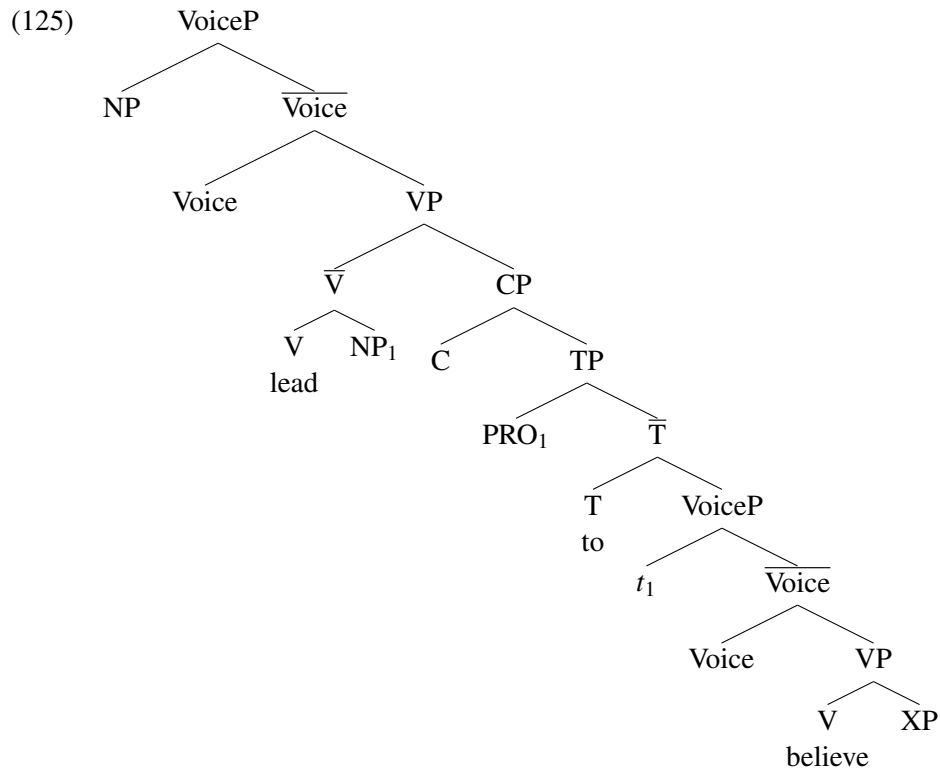
The Adv *seriously* selects a VP headed by *take*, and so they can form a conventionalized expression. The NP object can be open, because the conventionalized expression does not include any structurally superior NP.

The revised constraint also permits a collocation like *can't stand X*. Suppose this has a form something like the following (the exact details are not so important):



Here, NEG selects ModP, Mod select Voice, and Voice selects a VP headed by *stand*. Both the specifier of Voice and the complement of V can be open slots, because there is no fixed NP that is part of the collocation (see 117).

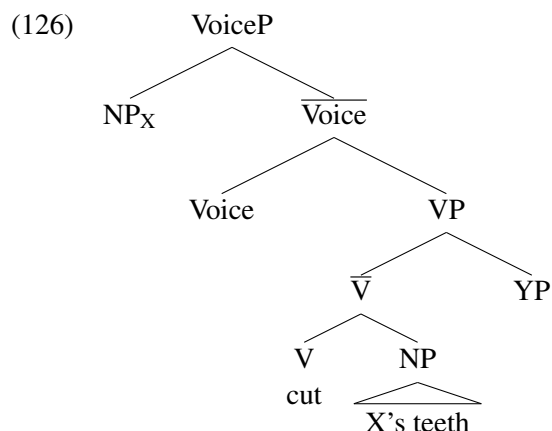
Consider also the collocation *lead X to believe Y*:



The verb *lead* selects a non-finite CP. Within the CP, C selects T, T selects Voice, Voice selects PRO and a VP headed by *believe*. If PRO is simply a variable of some kind, specified to co-vary with the higher object, then it is not truly a specified part of the conventionalized expression in the same way as a lexical NP. Then

there can be additional NPs within XP without violating the Constraint on Conventionalized Expressions (117).

Finally, there is also the idiom *X cut X's teeth Y-ing*. This one does have a fixed NP that precedes a slot that could itself be an NP (as a gerund), or it could include an NP within it (*Stephanopoulos cut his teeth quashing stories about Bill Clinton*). In order to not violate the Constraint on Conventionalized Expressions in (117), the fixed NP must not be structurally superior to the open Y position. This is in keeping with standard assumptions, since the NP *X's teeth* is a selected argument of the verb while the other phrase is an adjunct. It might therefore have this structure:



The set of nodes dominating *X's teeth* ($\{\bar{V}, VP, \overline{\text{Voice}}, \text{VoiceP}\}$) is not a proper subset of the set of nodes dominating YP ($VP, \bar{\text{Voice}}, \text{VoiceP}$) or anything within YP.

3.5 Summary

In section 2, we saw that collocations and idioms both seem to be subject to the same structural constraints. In section 3, the selection theory of idioms was extended to cover all *conventionalized expressions*, both literal collocations and non-literal idioms. All conventionalized expressions consist of local selectional relations between items, as codified in the Principle of Conventionalization (103). The Constraint on Conventionalized Expressions (117) permits open NP slots in low positions, so long as there is no fixed NP higher in the conventionalized expression. This rules out the missing patterns, *NP V X and *V NP X.

4 On Analyses of Ditransitives

In this section, I show that conventionalized expressions can be used to decide between competing syntactic analyses. The phenomenon I examine is ditransitives, already discussed to some extent in section 3. Here I show that only the ApplP analysis of Bruening (2010) succeeds in accounting for the patterns that we find. The previous section showed how the ApplP analysis works, and showed that it does indeed account for the patterns of conventionalized expressions. This section goes through other recent proposals for the structure of ditransitives, and shows that they all stumble in accounting for the patterns.

4.1 Larson 2017

I begin with a recent work, Larson 2017 (some of the content of which also appeared in Larson 2014). According to Larson (2017), idioms tell us nothing about analyses of ditransitives. There are no relevant ditransitive idioms according to Larson. Some phrases that have been treated as idioms, like *give X the creeps*,

are actually collocations. According to Larson, the noun *creeps* has a meaning ('a feeling of revulsion or fear'), which combines compositionally with the literal meaning of *give*. The reason that *creeps* does not occur with very many other verbs besides *give* and *get*, according to Larson, is that *give X the creeps* and *X get the creeps* are collocations. There is nothing more to explain, since collocations are not subject to any structural constraints, in Larson's view.

The study reported here shows that this is incorrect. Collocations are subject to the same structural constraints as idioms. Furthermore, as we have now seen, there are numerous collocations of the form V X NP, with the verb and second object part of the collocation but the first object open (*do X good/harm, give X permission, give X a break,...*). If Larson were correct and nouns with particular (sometimes idiomatic) meanings can just combine with verbs as collocations, with no structural constraints, we should also expect there to be collocations of the form V NP X, with the first NP fixed and the second one open. There are plenty of NPs with special meanings analogous to *the creeps* but which are animate, and should therefore be able to combine with *give* or another verb as its first object. Here is a short list of some animate NPs that would fit the bill:

- (127) a clockwatcher, a clothes horse, a Dutch uncle, everybody and his uncle, jailbait, the little woman, a living soul (NPI), a love child, a mother hen, the old ball and chain, X's old man, poster child, sharkbait, shrinking violet, smartypants, a snake in the grass, snowbunnies, star-crossed lovers, a wallflower, a wannabe

On Larson's reasoning, we should expect some of these to form collocations with verbs, perhaps like the following:

- (128) a. * give everybody and his uncle X
 b. * give the smartypants X
 c. * give X's old man Y
 d. * deny the wannabe X
 e. * give the little woman X
 f. * ask the/a shrinking violet X
 g. * ask the/a wallflower X

However, no collocations of this form exist, and it would be completely unnatural to conventionalize one of these.

Additionally, the first object of a double object construction can be inanimate in light verb constructions like the following:

- (129) give the table a scrubbing, give the house a new coat of paint, give the car a tuneup, give it a kick

There then ought to be collocations consisting of V NP_{inan} X, where X is some sort of eventive NP that forms light verb constructions. No such collocations exist, they are all V X NP instead:

- (130) V X NP:
 give X the once-over, give X (e.g. rugs) a beating, give X a try, give X a whirl

- (131) *V NP X:
 *give the rugs X, *give the car X, *give the bed X

Contra Larson (2017), it is truly significant that there are no idioms or collocations of the form *V NP X. This pattern is strikingly missing from ditransitives, in contrast with V X NP, which is well represented with

both collocations and idioms. There must be a syntactic constraint against conventionalized expressions of this form, since there is no other reason they could not be formed. The Constraint on Conventionalized Expressions in (117) explains their absence, together with the ApplP analysis of double object structures.

Larson (2017) further argues that in all the collocations and idioms with *give* of the form V NP to X, like *give way to X*, *give rise to X*, *give* plus the noun after it form an inseparable compound and do not constitute a phrasal idiom. His arguments for this are that: (i) they cannot passivize; (ii) the noun cannot be modified; (iii) the noun cannot be questioned. None of these arguments is telling, and some are incorrect. At least one of them can passivize (132a), and some of them can be modified (132b):

- (132) a. In build and coat and brush he was a huge timber-wolf; but **the lie was given to** his wolf-hood by his color and marking. (Jack London, *Brown Wolf*)
b. ...his courage **gave immediate way to** despair; (James Pettit Andrews, *History of Great Britain from the death of Henry VIII to the accession of James VI of Scotland to the crown of England*)

As for questioning, many if not most idioms ban questioning of their parts:

- (133) a. * What did she kick? The bucket.
b. * What did he screw? The pooch.
c. * What can't he hold to his father? A candle.

This is because many sub-parts of idioms do not have independent meaning.

Larson's arguments regarding V NP to X idioms like *give way to* do not go through, then. I see no reason to treat these idioms as being in any way different from other idioms, many of which also cannot passivize (e.g., *kick the bucket*).

This gives us the following patterns of conventionalized expressions in ditransitives (listed as "idioms; collocations," if that distinction is clear):

- (134) *double object construction*
a. V X NP: cry X a river, cut X some slack, throw X a bone; give X a break, give X a go
b. * V NP X: *throw the wolves X, *feed the lions X; *give X's old man Y
c. V NP NP: give the devil his due, give it a rest, give it the old college try
- (135) *prepositional dative construction*
a. V NP to X: give the lie to X, give rise to X; give priority to X, give some thought to X
b. V X to NP: feed X to the lions, throw X to the wolves;
c. V NP to NP: give hostages to fortune;

According to Larson (2017), it is important that we only look at prepositional datives that have a caused possession rather than a caused location meaning. This is because only these prepositional datives have the potential to alternate with the double object construction, which uniformly encodes caused possession. I list in (135) only those collocations and idioms that do indeed seem to have a caused possession meaning. The tests that are supposed to distinguish caused possession from locative PPs all show that these are caused possession. One such test involves substituting different directional prepositions for *to*. For instance, with *throw*, one can throw something to, at, toward, under, or over something. According to Hallman (2015), this test shows that *throw to the wolves* is not caused possession, because *throw* can occur with other prepositions. However, I would contend that, because this idiom itself does not permit any other preposition, this test actually shows that this idiom does encode caused possession. Intuitively, the subject of *throw* in the idiom does have the intent of giving the object to the wolves (intended caused possession). Turning to other idioms,

the preposition substitution test does show clearly that *give hostages to fortune* and *feed X to the lions* are caused possession: *give hostages to/*at/*toward*; *feed X to/*at/*toward*.

Another test for caused location versus caused possession is the ability to add *halfway* (Rappaport Hovav & Levin 2008): *threw it halfway to the goal line* is wellformed, showing that *throw* may involve location and not caused possession. But *halfway* is not possible with the idiom *throw X to the wolves*: **They threw him halfway to the wolves*. This test also shows that *give hostages to fortune* and *feed X to the lions* are caused possession: **gave it halfway to Bill*, **fed it halfway to the dog*.

A third test is the ability to question the PP with *where* rather than *to who(m)*. *Where did you throw it?* is wellformed, showing that *throw* is (or can be) a location verb (but again, this does not work with the idiom *throw X to the wolves*). This test also shows that *give hostages to fortune* and *feed X to the lions* are caused possession: **Where did you give it?* **Where did you feed it?* (where *it* is the food item).

In addition, according to Rappaport Hovav & Levin (2008) and Larson (2014), the verbs *give*, *hand*, and *lend* can only ever encode caused possession, so any conventionalized expression with those verbs is relevant (which makes *lend X color*, *Y lend X Y's ear*, *lend X a hand* all examples of the V X NP pattern).

This means that, contra Larson (2017), there are in fact many conventionalized expressions that are caused possession ditransitives. Among prepositional datives, all possible combinations of the verb and at least one of its objects are attested (135). In contrast, among double object constructions, one of the three possible patterns is systematically missing (see 134b). Given that the pattern with both objects exists (134c), it cannot be the case that something rules out the first object being part of a conventionalized expression. It is only impossible for the first object to be part of the expression if the second is not. The Constraint on Conventionalized Expressions in (117) explains this curious fact, in conjunction with the ApplP analysis of double object constructions. This was shown in section 3.3.3.

As just observed, all possible patterns of the prepositional dative exist. In particular, both V NP P X (*give rise to X*) and V X P NP (*feed X to the lions*) exist. The former regularly alternates with the double object construction, as in *give the lie to X* alternating with *give X the lie*. Significantly, the second pattern never does: *throw X to the wolves* and *feed X to the lions* do not alternate with **throw the wolves X* or **feed the lions X* (Rappaport Hovav & Levin 2008, Bruening 2010, Larson 2014). This is a systematic pattern that needs an explanation. The Constraint on Conventionalized Expressions in (117) does explain it, in conjunction with the ApplP analysis of ditransitives, as was shown in section 3.3.3.

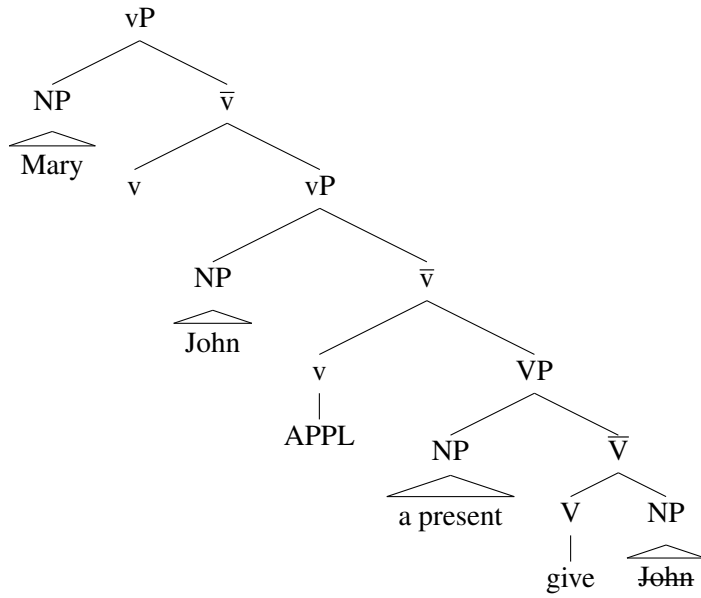
The important point here is that there are systematic patterns of conventionalized expressions in ditransitives, contra Larson (2017). They can and should be used to decide between analyses of ditransitive constructions. Having concluded this, I will go on in the next subsections to show that the patterns rule out various analyses that have been proposed for double object constructions. Only the ApplP analysis discussed in section 3.3.3 succeeds in capturing the patterns depicted in (134–135).

4.2 Larson 2014

Larson (2014: 102) proposes that the double object construction has the following structure, where the first object starts as the complement to the verb *give* and then moves to the specifier of a higher vP (which Larson identifies as Appl):⁵

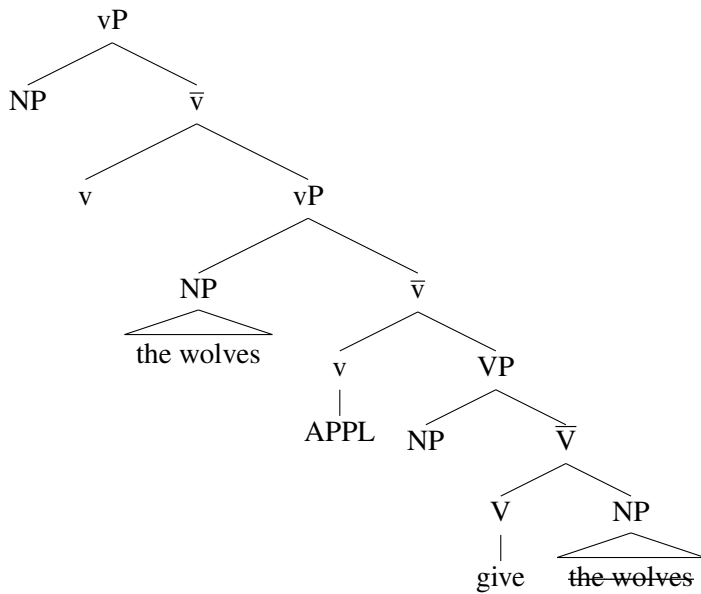
⁵This tree is the slightly simplified version given in Larson (2017: 395, (12b)). It differs from the tree shown in Larson (2014: 102, (177b)) only in that it does not depict details of the analysis like syntactic features and Agree relations.

(136)



This analysis incorrectly predicts that the missing pattern of conventionalized expressions should exist. Take a non-existent conventionalized expression like **give the wolves X*. This would have the following structure:

(137)



This would satisfy every proposed constraint on idioms: *give* and *the wolves* form an underlying constituent, excluding all idiomatic material; and the verb *give* selects the object *the wolves*. Adopting the Constraint on Conventionalized Expressions in (117) does not help: the set of nodes that dominate *the wolves* is not a proper subset of the set of nodes that dominate the open NP (following Larson 2017 and every other work on this topic, only the underlying positions of syntactic constituents should matter to idioms). Larson's analysis has no way to rule out this missing pattern.

4.3 Other Derivational Analyses

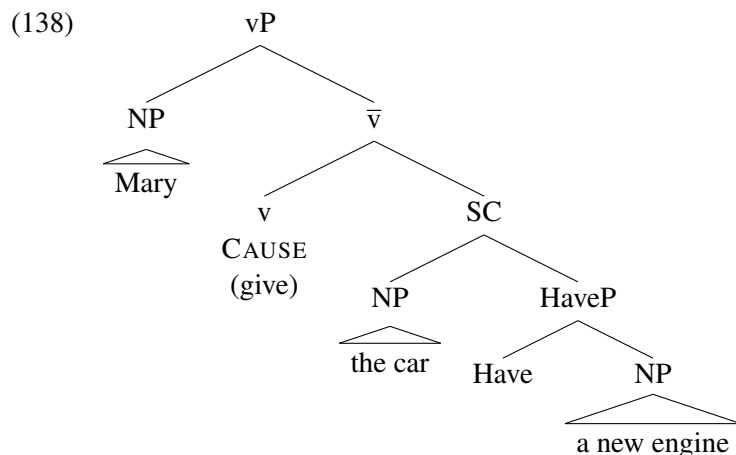
Larson's 2014 analysis is based on the claim that the order of the prepositional dative is basic, and the double object construction is derived from it (put another way, the recipient starts out lower than the theme

and moves across it). There have been other proposals to this effect, starting with Chomsky (1955/1975). The analysis in Larson (1988) is basically the same as the analysis in Larson (2014). Baker (1988) proposed a preposition incorporation analysis. So did Ormazabal & Romero (2010). Baker (1996: 20, (36)) proposed a derivational analysis that is almost identical to that of Larson (2014). In the variant proposed by Aoun & Li (1993: 35, (62)), the first object starts below the second and incorporates into the verb, moving with it to a higher position. In the variant proposed by den Dikken (1995: 132, (52b)), the first object starts in a small clause below the second, and moves across the second object to the specifier position of a higher small clause. Oba (2002: 43, (13)) proposes something similar. All such analyses suffer from the same problem as Larson's: they incorrectly predict that the missing *V NP X pattern should exist, since the V X to NP pattern does. They cannot appeal to a constraint banning the transformation just when the recipient is part of a conventionalized expression, because it is possible to have both objects be part of one, as in *give the devil his due*, *give it your best shot*. It is also possible for the V X NP and V NP to X patterns to alternate with each other, as was shown above (*read X the riot act*, *read the riot act to X*); there is therefore nothing wrong in principle with conventionalized expressions participating in the proposed transformation.

For further reasons to reject analyses that derive double object constructions (or applicative constructions more generally) from oblique PP frames, see Bruening (2018b).

4.4 HaveP (Harley 1997 etc.)

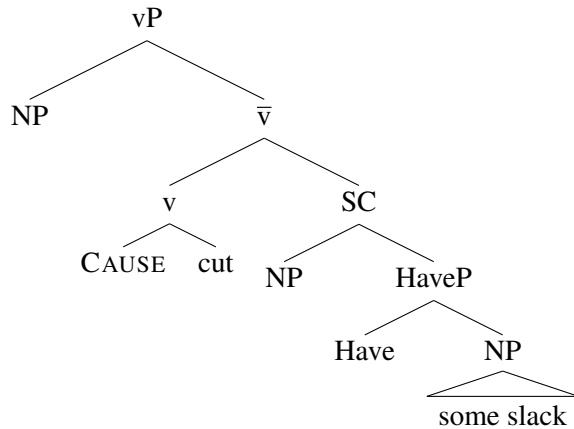
A currently prominent analysis of double object constructions proposes that they embed a Have component under a Cause component. This is the HaveP analysis of Harley (1997, 2002, 2008), Beck & Johnson (2004), Harley & Jung (2015), and others; a variant of it also appears in Ramchand (2008: 103). The structure that is proposed is the following, where a causative *v* head takes a small clause as its complement ("SC"), headed by Have. Have moves to *v*, where the two together are pronounced *give*.



This analysis can explain the missing *V NP X pattern (**throw the wolves X*). This pattern violates the Constraint on Conventionalized Expressions in (117). A structurally superior NP is included in the idiom (Spec-HaveP) while a structurally lower NP is not (complement of Have). The set of nodes that dominate the subject of the small clause is a proper subset of the set of nodes that dominate the complement of Have.

Unfortunately, this analysis runs into problems with conventionalized expressions with verbs other than *give*. According to Harley (2008), verbs other than *give* have that verb stem adjoined to *v* as a “manner modifier”:

(139)



The above structure would have to be the analysis of the idiom *cut X some slack*. The problem is that there is no relation between the verb stem *cut* and the NP *some slack*. The NP is certainly not selected by the verb stem in the above structure, instead it is selected by *Have*. *Cut* and *slack* also do not form an underlying constituent. There is therefore no way for the two to be interpreted as an idiom, on any theory of idioms.

This problem is even more severe for idioms that occur as simple transitives, but can be used as double object constructions. The following examples are canonically just a verb and an object, but they can also be used as double object constructions with an open first object:

- (140) a. pull strings — pull X some strings
b. do the honors — do X the honors (e.g., ‘do me the honors’)

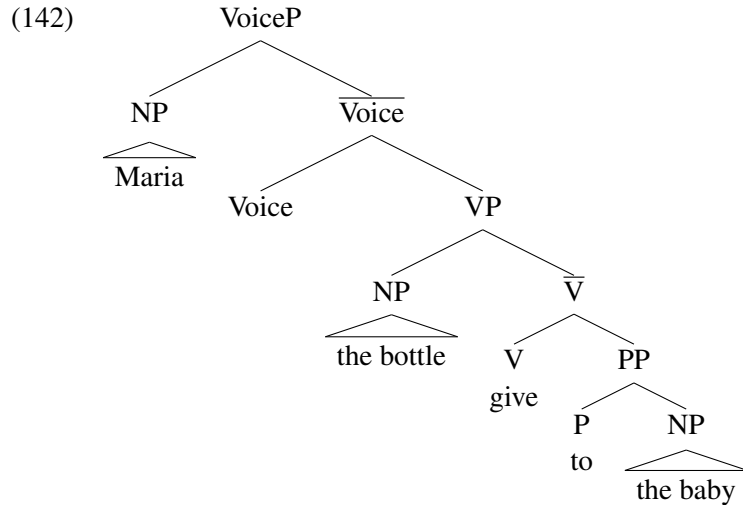
In the case of these idioms, we clearly want to say the idiom is just the verb and its object. It does not involve *Have* at all. Somehow the idiomatic interpretation would have to be preserved when the verb stem is adjoined to *v* as a manner modifier, and the NP is selected by *Have* rather than by the verb stem. Permitting such a discontinuous idiom would incorrectly allow all kinds of discontinuous idioms.

The *HaveP* analysis also runs into problems with conventionalized expressions that can alternate between the double object construction and the prepositional dative construction. The following examples have an open object of a preposition in the canonical variant, but can also appear in the double object construction:

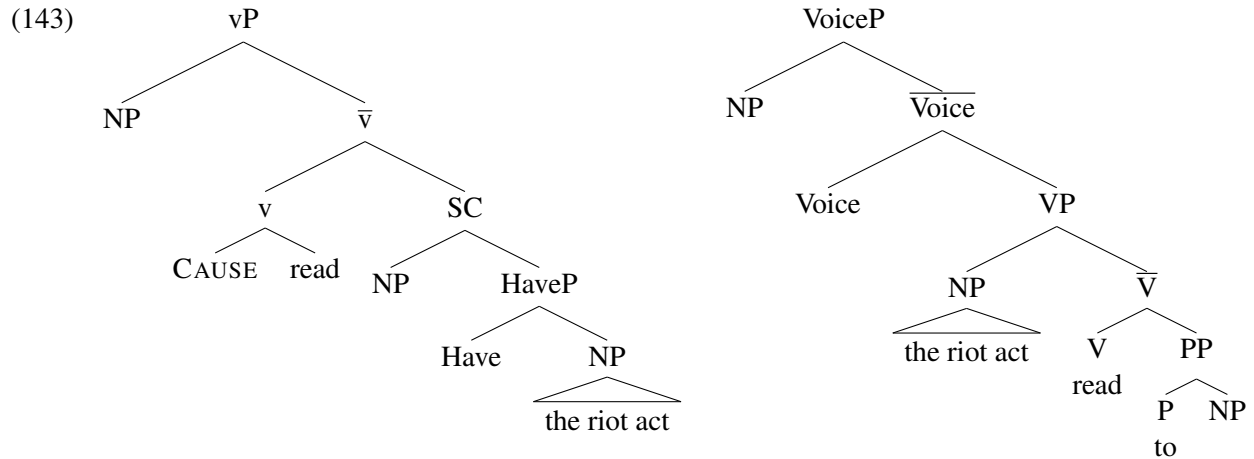
- (141) a. make allowances for X — make X allowances
b. open (some/a few) doors (for X) — open X some doors
c. pass the torch (on) to X — pass X the torch
d. pay court to X — pay X court

There are also idioms and other conventionalized expressions that freely alternate between the two, for instance *read the riot act to X — read X the riot act* (Bruening 2010, Larson 2017), and the collocations in (30).

Harley & Jung (2015: 714, (19)) adopt the following structure for the prepositional dative construction:



The problem with alternating conventionalized expressions is that, in the HaveP analysis, there is no shared piece of structure in the two constructions:



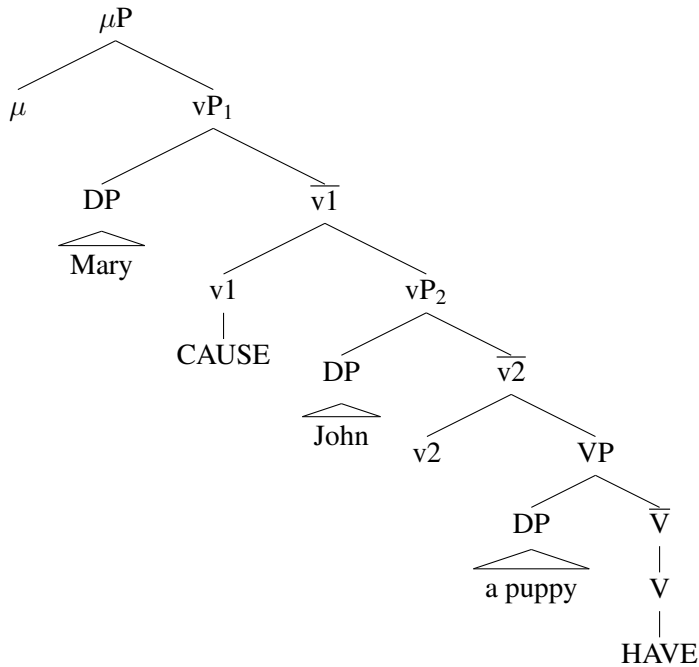
In the prepositional dative on the right, the NP is a selected argument of the verb *read* (but the Constraint on Conventionalized Expressions is violated), but in the double object construction on the left it is instead a selected argument of *Have*, and the verb stem bears no relation to the NP whatsoever.

I conclude that the HaveP analysis is incapable of accounting for the patterns of conventionalized expressions that we find. See also Bruening (2018a) for further reasons to reject the HaveP analysis of double object constructions.

4.5 Hallman 2015

Hallman (2015) adopts a different version of the HaveP analysis (Hallman 2015: 394, (12)):

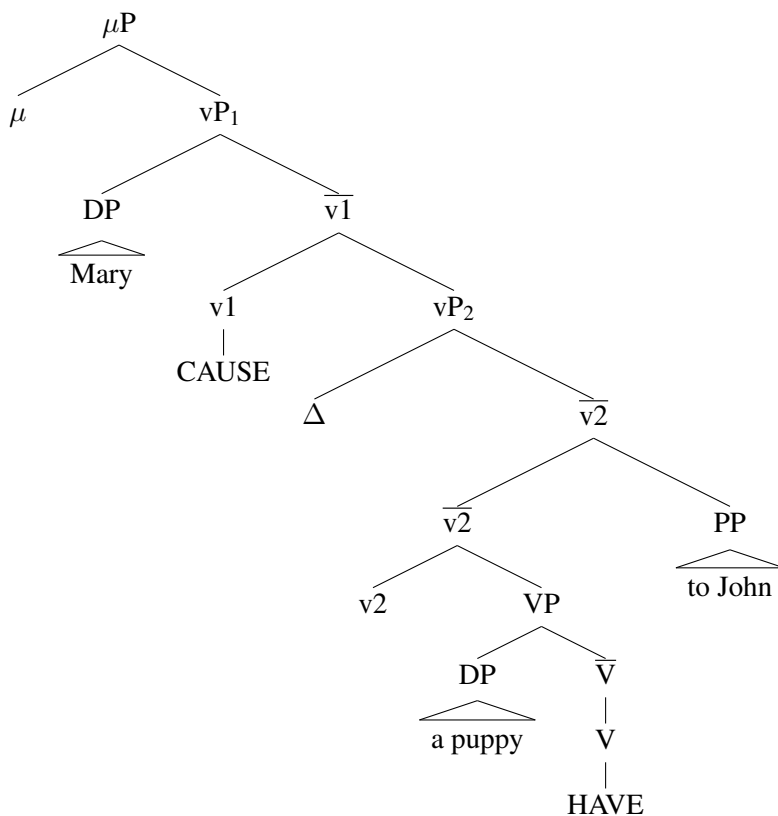
(144)



The second object is the specifier of a lower V HAVE, while the first object is the specifier of a higher v. A CAUSE v takes that as its complement, and projects the external argument. Once again, HAVE moves up to CAUSE, where it is pronounced *give*.

Hallman proposes that the prepositional dative construction with caused possession verbs like *give* is derived from the above structure by demotion of the specifier of $v2$ to adjunct status, with movement of the lower object (*a puppy*) to the now-open Spec- $v2$, marked with Δ (Hallman 2015: 400, (21)):

(145)



The problem with this analysis is that it cannot rule out the missing *V NP X pattern (**throw the wolves X*) without also ruling out the attested V X to NP pattern (*throw X to the wolves*). Hallman claims that idioms like *throw X to the wolves* are not caused possession prepositional datives, instead they are locative PPs. Locative PPs have a different structure, and are not derived from the double object structure. He claims that no caused possession verb could ever form a V X to NP idiom, just as they could never form the missing *V NP X pattern.

This is not correct, however. There are caused possession V NP/X to NP idioms, as we saw in section 4.1. Two examples are *give hostages to fortune* and *feed X to the lions*. As we saw in section 4.1, the tests that are supposed to distinguish caused possession from locative PPs all show that these idioms involve caused possession. It is therefore clear that Hallman is incorrect that caused possession idioms do not occur in the prepositional dative. They do. Yet they do not alternate with the double object construction. This is inexplicable in Hallman's transformational analysis, where caused possession double object constructions and caused possession prepositional dative constructions have the same underlying source.⁶

Hallman (2015) also claims that there must be a syntactic difference between different types of PPs, in particular caused possession PPs with verbs like *give* and locational PPs with verbs like *put*. In the analysis proposed in section 3.3.3 and adopted throughout this paper, all PPs are treated as having the same structure. Any differences come only from the lexical semantics of the verb, not from a structural difference. Hallman (2015) claims that there is a syntactic difference between caused possession PPs and locational PPs in their ability to control into non-subject-gap purpose clauses, but this is false. Locational PPs can control into purpose clauses, contra Hallman (thanks to Larson Stromdahl for discussion and examples):

- (146) a. I put it beside her₁ [PRO₁ to look over when she got the chance].
 b. I set it in front of her₁ [PRO₁ to sign].
 c. (What did you do with the microprocessor?) We inserted it in him₁ [PRO₁ to track himself₁ with].

Hallman's examples all seem to be unacceptable for pragmatic reasons.

Hallman also claims that the purported difference in the ability to control is due to a difference in c-command, such that the object of a caused possession P c-commands the purpose clause but the object of a locational P does not. However, c-command is not required at all for non-subject-gap purpose clauses, as Whelpton (2002) showed. For instance, the possessor of the object of a locational P can also control into a purpose clause (these examples are also from Larson Stromdahl):

- (147) a. I put it in her₁ mailbox [PRO₁ to look over when she returned from the Bahamas].
 b. I threw it into his₁ cell [PRO₁ to eat after his trial].

Hallman's claimed differences between the two types of PPs have no empirical basis, then. As far as I can tell, there are no syntactic differences between caused possession PPs and other argument PPs like locations. I therefore analyze them all as having the same structure, that proposed in section 3.3.3.

As we have seen, Hallman's 2015 analysis is built on a misunderstanding of the facts. Caused possession PPs do not differ structurally from locational PPs, and there are caused possession idioms. Hallman's analysis cannot account for them.

⁶It is true that caused possession idioms like *feed X to the lions* are relatively rare. However, this just follows from the general preference for fixed NPs in idioms to be inanimate (Nunberg et al. 1994). The possessor in a caused possession construction generally has to be animate. Since there are at least two idioms of this form, there could not be a structural constraint as Hallman claims.

4.6 Other Analyses of Ditransitives

Other proposals that I will not address in any detail also make the wrong predictions. For instance, Kayne (1984) proposed that double object constructions involve a small clause. In this analysis, neither NP is an argument of the verb, instead the first object is the subject of a small clause whose predicate is the second object. This analysis cannot account for the fact that the *V NP X pattern is missing while the V X NP pattern is robustly attested. Neither NP is selected by the verb in the small clause analysis, and neither forms a constituent with it to the exclusion of the other. The two NPs also c-command each other, making the Constraint on Conventionalized Expressions (117) irrelevant. (Note that in the versions of the small clause analysis proposed by Johnson (1991) and Hornstein (1995), the first NP does c-command the second, so these versions could potentially account for the patterns of conventionalized expressions by adopting the Constraint on Conventionalized Expressions. However, see Bruening (2018a) for other reasons to reject small clause analyses.)

Pesetsky (1995: 197, (511)) proposes a structure where the recipient in the double object construction starts out in a structurally higher position than the theme, with a null preposition that Pesetsky calls “G” introducing the theme. This structure can capture the double object patterns, since the verb can select the P and the P selects the theme; the missing V NP X pattern can be ruled out by adopting the Constraint on Conventionalized Expressions in (117). The problem for Pesetsky (1995) is the prepositional dative construction. In Pesetsky’s analysis, this has the same underlying hierarchical structure, with the theme moving to a position above the goal (223, (565b)). This incorrectly rules out the attested V X to NP pattern (*throw X to the wolves*), because it would violate the Constraint on Conventionalized Expressions (117) for the same reason as the double object construction.

Pylkkänen’s 2008 “low applicative” analysis has the potential to explain the asymmetry in the double object construction. In her analysis, the V takes an ApplP complement, with the first object in the specifier of Appl and the second object the complement of Appl. Adopting the Constraint on Conventionalized Expressions in (117) would explain the missing *V NP X pattern. Pylkkänen (2008) does not give a structure for the prepositional dative, however; assuming that it does not have an ApplP projection, her analysis would run afoul of the same problem as the HaveP analysis (section 4.4). There would be no piece of structure in common between the prepositional dative and the double object construction. Alternating conventionalized expressions could not be accounted for: in *read the riot act to X*, the verb would select *the riot act*, but in *read X the riot act*, Appl would instead. (See also Larson 2010 on problems with Pylkkänen’s proposed semantics.)

4.7 Summary

The ApplP analysis of double object constructions succeeds in accounting for all of the patterns of conventionalized expressions that we find in ditransitives, as was shown in section 3. In this section, I showed that no other proposal for the structure of ditransitives can account for all of the patterns. Conventionalized expressions, then, successfully distinguish between analyses of certain empirical phenomena.

5 Conclusion

The detailed empirical study of collocations and idioms undertaken here has shown that they consist of almost identical syntactic structures. Contra Larson (2017), idioms and collocations behave the same. There are structural constraints on both idioms and collocations, and they are the same constraints. In particular, two patterns are systematically missing from both collocations and idioms: a pattern with a fixed subject and open object, and the missing *V NP X double object pattern. Also missing are discontinuities in selection, where two elements form a conventionalized expression but neither locally selects the other. The selection

theory of idioms proposed by Bruening (2010) extends to literal collocations as a theory of conventionalized expressions in general. A constraint against having an open NP argument slot if there is a structurally superior fixed NP rules out the two missing patterns.

Importantly, significant asymmetries and systematically missing patterns in conventionalized expressions can inform us about the proper analysis of syntactic phenomena. Again contra Larson (2017), conventionalized expressions are very telling regarding ditransitives. As was shown here, only the ApplP analysis of ditransitives (Bruening 2010) captures the facts. All other analyses of ditransitives fail to account for the patterns of conventionalized expressions.

Bibliography

- Aoun, Joseph & Yen-Hui Audrey Li. 1993. *Syntax of scope*. Cambridge, MA: MIT Press.
- Baker, Mark. 1988. Theta theory and the syntax of applicatives in Chichewa. *Natural Language and Linguistic Theory* 6. 353–389.
- Baker, Mark C. 1996. On the structural positions of themes and goals. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 7–34. Dordrecht: Kluwer.
- Baltin, Mark R. 1989. Heads and projections. In Mark R. Baltin & Anthony S. Kroch (eds.), *Alternative conceptions of phrase structure*, 1–16. Chicago: The University of Chicago Press.
- Barss, Andrew & Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17. 347–354.
- Beck, Sigrid & Kyle Johnson. 2004. Double objects again. *Linguistic Inquiry* 35. 97–123.
- Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and θ -theory. *Natural Language and Linguistic Theory* 6. 291–352.
- Bruening, Benjamin. 2001. QR obeys superiority: Frozen scope and ACD. *Linguistic Inquiry* 32. 233–273.
- Bruening, Benjamin. 2010. Ditransitive asymmetries and a theory of idiom formation. *Linguistic Inquiry* 41. 519–562.
- Bruening, Benjamin. 2013. By-phrases in passives and nominals. *Syntax* 16. 1–41.
- Bruening, Benjamin. 2014. Precede-and-command revisited. *Language* 90. 342–388.
- Bruening, Benjamin. 2018a. Depictive secondary predicates and small clause approaches to argument structure. *Linguistic Inquiry* 49. to appear.
- Bruening, Benjamin. 2018b. Double object constructions and prepositional dative constructions are distinct: A reply to Ormazabal and Romero 2012. *Linguistic Inquiry* 49. 123–150.
- Bruening, Benjamin, Xuyen Dinh & Lan Kim. 2017. Selection, idioms, and the structure of nominal phrases with and without classifiers. *Glossa* to appear.
- Chomsky, Noam. 1955/1975. *The logical structure of linguistic theory*. New York: Plenum.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

- Collins, Chris & Paul M. Postal. 2014. *Classical NEG raising*. Cambridge, MA: MIT Press.
- den Dikken, Marcel. 1995. *Particles: On the syntax of verb-particle, triadic, and causative constructions*. Oxford: Oxford University Press.
- Ernst, Thomas. 1981. Grist for the linguistic mill: Idioms and “extra” adjectives. *Journal of Linguistic Research* 1. 51–68.
- Finer, Daniel L. 1985. The syntax of switch-reference. *Linguistic Inquiry* 16. 35–55.
- van Gestel, Frank. 1995. En bloc insertion. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 75–96. Hillsdale, NJ: Lawrence Erlbaum.
- Green, Georgia M. 1974. *Semantics and syntactic regularity*. Bloomington: Indiana University Press.
- Hallman, Peter. 2015. Syntactic neutralization in double object constructions. *Linguistic Inquiry* 46. 389–424.
- Harley, Heidi. 1997. If you have, you can give. In Brian Agbayani & Sze-Wing Tang (eds.), *Proceedings of the west coast conference on formal linguistics xv*, 193–207. Stanford, CA: CSLI Publications.
- Harley, Heidi. 2002. Possession and the double object construction. *Yearbook of Linguistic Variation* 2. 29–68.
- Harley, Heidi. 2008. The bipartite structure of verbs cross-linguistically, or, why Mary can’t exhibit John her paintings. In Thaís Cristófaró Silva & Heliana Mello (eds.), *Conferências do v congresso internacional da associação brasileira de lingüística*, 45–84. Belo Horizonte, Brazil: ABRALIN and FALE/UFMG.
- Harley, Heidi & Hyun Kyoung Jung. 2015. In support of the P_{HAVE} analysis of the double object construction. *Linguistic Inquiry* 46. 703–730.
- Harwood, William. 2015. Being progressive is just a phase: Celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language and Linguistic Theory* 33. 523–573.
- Hornstein, Norbert. 1995. *Logical form: From GB to Minimalism*. Cambridge, MA: Blackwell.
- Janke, Vikki & Ad Neeleman. 2012. Ascending and descending VPs in English. *Linguistic Inquiry* 43. 151–190.
- Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9. 577–636.
- Kayne, Richard. 1984. Unambiguous paths. In *Connectedness and binary branching*, 129–163. Dordrecht: Foris.
- Koopman, Hilda & Dominique Sportiche. 1991. The position of subjects. *Lingua* 85. 211–258.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In John Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.
- Larson, Richard K. 2010. On Pylkkänen’s semantics for low applicatives. *Linguistic Inquiry* 41. 701–704.
- Larson, Richard K. 2014. *On shell structure*. London: Routledge.

- Larson, Richard K. 2017. On “dative idioms” in English. *Linguistic Inquiry* 48. 389–426.
- Lebeaux, David. 2009. *Where does binding theory apply?* Cambridge, MA: MIT Press.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
- Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam A. Mchombo (ed.), *Theoretical aspects of bantu grammar*, 113–150. Stanford: CSLI Publications.
- Marantz, Alec. 1997. No escape from syntax: Don’t try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *Proceedings of the 21st annual penn linguistics colloquium* (Penn Working Papers in Linguistics 4:2), Philadelphia: University of Pennsylvania.
- McGlone, Matthew S., Sam Glucksberg & Cristina Cacciari. 1994. Semantic productivity and idiom comprehension. *Discourse Processes* 17. 167–190.
- Moon, Rosamund. 1998. *Fixed expressions and idioms in English*. Oxford: Clarendon Press.
- Nicolas, Tim. 1995. Semantics of idiom modification. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 233–252. Hillsdale, NJ: Lawrence Erlbaum.
- Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70. 491–538.
- Oba, Y. 2002. The double object construction and the extraction of the indirect object. *Linguistic Analysis* 32. 40–71.
- O’Grady, William. 1998. The syntax of idioms. *Natural Language and Linguistic Theory* 16. 279–312.
- Ormazabal, Javier & Juan Romero. 2010. The derivation of dative alternations. In Maia Duguine, Susana Huidobro & Nerea Madariaga (eds.), *Argument structure and syntactic relations: A cross-linguistic perspective*, 203–232. Amsterdam: John Benjamins.
- Osborne, Timothy, Michael Putnam & Thomas M. Gross. 2012. Catenae: Introducing a novel unit of syntactic analysis. *Syntax* 15. 354–396.
- Penka, Doris. 2012. Split scope of negative indefinites. *Language and Linguistics Compass* 6. 517–532.
- Pesetsky, David. 1992. Zero syntax: Volume 2. Ms., MIT.
- Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-driven phrase structure grammar*. Chicago: Chicago University Press.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.
- Ramchand, Gillian Catriona. 2008. *Verb meaning and the lexicon: A first-phase syntax*. Cambridge: Cambridge University Press.
- Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics* 44. 129–167.
- Richards, Norvin. 2001. An idiomatic argument for lexical decomposition. *Linguistic Inquiry* 32. 183–192.

Riehemann, Susanne Z. 2001. *A constructional approach to idioms and word formation*: Stanford University dissertation.

Schenk, André. 1995. The syntactic behavior of idioms. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 253–271. Hillsdale, NJ: Lawrence Erlbaum.

Svenonius, Peter. 2005. Extending the extension condition to discontinuous idioms. *Linguistic Variation Yearbook* 5. 227–263.

Whelpton, Matthew. 2002. Locality and control with infinitives of result. *Natural Language Semantics* 10. 167–210.

Department of Linguistics and Cognitive Science
University of Delaware
Newark, DE 19716
(302) 831-4096
bruening@udel.edu