

Idioms, Anaphora, and Movement Diagnostics

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

draft December 3, 2015, comments welcome

Abstract

Phrasal idioms are often used to argue that some syntactic phenomenon involves a movement dependency. This paper argues that the only requirement on idioms is that a chunk of an idiom not receive an interpretation that is inconsistent with its idiomatic interpretation. Once this constraint is satisfied, idiom chunks can participate in pronominal anaphora and control in addition to movement. They can also appear in thematic argument positions separate from their idiom, even separated from the rest of their idiom by sentence boundaries. A detailed investigation of various dependencies including raising, control, pronominal anaphora, *do so* anaphora, parasitic gaps, relative clauses, and tough-movement shows that idioms are compatible with analyses that do not involve direct movement (e.g., null operator analyses). This paper proposes a theory of idioms in which all lexical entries are partial syntactic structures paired with phonological items. Idiomatic readings are triggered when part of a syntactic structure being processed matches the stored partial structure. This theory correctly accounts for all of the facts of anaphora and movement involving idioms.

1 Introduction

Since at least an unpublished 1968 manuscript by Brame (cited in Schachter 1973 and Vergnaud 1974), the syntactic literature has often appealed to phrasal idioms to motivate certain analyses of syntactic phenomena. In particular, idioms are often used to justify a *movement* analysis of some phenomenon. For instance, Brame's own examples involved relative clauses, and recent years have seen a revival of the head raising analysis of relative clauses that he first proposed (e.g., Kayne 1994, Bianchi 1999, Hulsey and Sauerland 2006, Henderson 2007). It is worth quoting some of this literature, concerning examples like that in (1) (where the idiom is *make headway*; in all examples the parts of the idiom will be italicized):

- (1) The *headway* that John *made* was impressive.

The reasoning goes like this: "Since verb-complement idioms must be interpreted as a unit, it must be the case that the external head in [(1)] is interpreted in the complement position within the [relative] clause" (Henderson 2007, 203). Thus, the head of the relative clause, *headway*, must have started within the relative clause, as complement to *made*, and moved to its surface position. Similarly: "If the idiomatic reading of *headway* is available only if *headway* appears inside the complement of *make*, [(1)] requires the raising analysis" (Hulsey and Sauerland 2006, 113).

The raising analysis of relative clauses should be contrasted with another possible analysis, one where a (null or overt) relative operator moves, while the head noun itself never occupies a position within the relative clause (e.g., Chomsky 1977). Idioms, this argument goes, show that such an analysis is untenable.

This reasoning is not limited to relative clauses. It has played a very important role in the literature on raising versus control (e.g., Postal 1974). As one recent example of many, Postal (2004, 93) states that certain idioms may only undergo raising and do not participate in control, and so they can be used to identify movement phenomena. Some of Postal's examples of this contrast are the following:

- (2) (Postal 2004, 93, 26b–c)
- a. They believe *the shit* to have *hit the fan* yesterday. (raising to object = movement)

- b. *They convinced *the shit* to *hit the fan* yesterday. (control; no movement)

Postal then uses idioms as a diagnostic to argue for a process of raising to object of a preposition.

At the same time, other researchers have pointed out numerous issues for the use of idioms as a movement diagnostic, for instance Ruwet (1991) and Nunberg, Sag, and Wasow (1994). They have noted that idioms can, in fact, participate in control (3). They can also participate in pronominal anaphora and VP ellipsis (4):

(3) *Control:*

- a. To some, the lesson is that no matter how happy the tune may be, eventually, *the piper* wants to be *paid*. (Riehemann 2001, 62, (60a))
- b. *The plug* is waiting to be *pulled* once all the 5600 stock is gone...
(<http://www.oesf.org/forum/lofiversion/index.php/t1554.html>)
- c. ... but budget constraints forced *corners* to be *cut* elsewhere, resulting in a 5atm watch.
(<http://www.christopherwardforum.com/viewtopic.php?f=1&t=33473&start=30>)

(4) *Pronominal Anaphora*

- a. We thought *the bottom* would *fall out of* the housing market, but **it** didn't. (Nunberg, Sag, and Wasow 1994, (7b))
- b. A: *The shit* will *hit the fan* tonight! B: Yes, **it** will.

Such examples call into question the utility of idioms as a movement diagnostic, as chunks of idioms can apparently be related to the rest of their idiom by syntactic relations other than movement.

In this paper, I explore such data in detail and show that idioms can participate in numerous syntactic dependencies other than movement. Rather than idioms being a diagnostic of movement, I argue that the following constraint holds:

(5) The Constraint on Idiom Chunks (CIC):

Where X is a sub-part of an idiomatic expression Y and X appears in position 1 but is related to another position, position 2, through some anaphoric dependency, X must not receive an interpretation either in position 1 or position 2 that is inconsistent with its interpretation in Y.

The CIC can be viewed as following from something like the Anti-Pun Ordinance of Zaenen and Karttunen (1984, 316), which bans elements from having two different meanings simultaneously.

The way the CIC explains the raising/control contrast in (2) is that *the shit* does not receive any interpretation in its derived object position in (2a). Semantically, raising to object verbs simply take a proposition as their complement. Example (2a) is semantically identical to *They believe that the shit hit the fan yesterday*. In contrast, in the object control case in (2b), *the shit* is being assigned an interpretation in the higher object position: it is the object of *convince*, which requires a sentient entity that can be swayed by arguments. This interpretation is incompatible with the meaning of *the shit* in the idiom, which is not sentient, and the CIC is violated. In the examples of control in (3), in contrast, *want*, *wait*, and *force* do not necessarily attribute sentience or feelings to the NP that serves as their argument. For instance, *want* as a subject control verb can occur with inanimates and weather-it as its subject:

- (6) a. The stone wants to be stone; the artist wants it to be art.
b. I applied a coat of Ultra Grip, as sometimes metal doesn't want to be painted.
c. Some days it feels like it wants to rain.

Want appears to be able to refer to a situation or properties of the world, without attributing sentience or desires to its subject in particular. As such, no interpretation is being assigned to *the piper* in (3a) that is inconsistent with the idiom. *Wait* in (3b) behaves similarly to *want* (*it's waiting to rain; the stone waits to receive the sculptor's impressions*). In (3c), *force* can refer to the bringing about of a situation without direct action on the NP that acts as the syntactic direct object of *force* (e.g., *Those changes forced the bill to be sent back to the House one more time*).

These properties of these control verbs are explored more fully in section 4. Finally, in the examples of pronominal anaphora in (4), both the idiom chunks and the pronouns that are covalued with them occur only in the context of the idiom (mediated through VP ellipsis, elaborated on below). This is also consistent with the CIC.

Note that part of the justification for the view that idioms are a valid movement diagnostic is the idea that idiom chunks could not possibly occupy any thematic argument position other than the one within their own idiom. The claim is that they can only occupy other positions if those positions are non-thematic, hence, potentially empty positions into which movement is possible. The CIC, in contrast, imposes no such restriction. As the examples of control in (3) show, the CIC is upheld, and the movement view is not.

In the rest of this paper, I go through various different syntactic dependencies and show that in every case, the CIC makes the correct predictions, while viewing movement as the only dependency that is possible for an idiom chunk fails. Section 3 discusses pronominal anaphora and VP ellipsis; section 4 goes through raising and control in detail; and section 5 discusses idioms in parasitic gap constructions. In every case, we will see that idiom chunks are constrained only by the CIC, and are not limited to movement dependencies. Idiom chunks can participate in control and they can antecede pronouns. If parasitic gaps involve null operators (as in Chomsky 1986), they can also form relations with null operators.

Section 6 then turns to two contentious cases where idioms have been used to argue for a movement analysis over an operator analysis, namely, tough-movement and relative clauses. I show there that idioms are consistent both with direct movement analyses and with (null) operator analyses. I also discuss relative clauses in more detail, as they raise particular issues for all analyses, including the CIC. I show that there are a range of interpretations involved in relative clauses with idiom chunks as heads, none of which actually violates the CIC.

Finally, section 7 outlines a theory of idioms that predicts exactly the range of facts presented here. I argue that idioms are stored pairings between syntactic structures and phonological items (as are words generally). Idiomatic readings are activated when a sub-part of a syntactic structure being processed matches the stored pairing. An anaphoric element as a sub-part of such a structure can also lead to matching, since anaphoric elements trigger the activation of the lexical content of their antecedent. I show that this theory predicts the existence of some unexpected cases of cross-sentential anaphora with idiom chunks, cases which are problematic for every other theory of idioms.

Before turning to syntactic dependencies, I first outline the empirical domain of inquiry in section 2.

2 Preliminaries: Idioms and Other Fixed Expressions

I begin with some preliminary remarks on how to identify idioms in order to delimit the domain of inquiry in this paper. I consider idioms to be a sub-type of fixed expression, defined below:

- (7) Fixed Expression: A phrase where the particular words used are conventionally fixed and cannot be varied.

Some examples of fixed expressions that are interpreted entirely literally appear below:

- (8) a. No can do.
b. Til death do us part.
c. give X a round of applause
d. a concerted effort
e. last but not least

Phrasal idioms are then fixed expressions where at least a sub-part of the expression is interpreted in a non-literal manner. Moreover, it is crucial that this non-literal interpretation arises only in the context of the rest of the fixed expression.

- (9) Phrasal Idiom: A fixed expression a sub-part of which is interpreted in a non-literal way just in the context of that fixed expression.

I illustrate with three examples, only two of which are phrasal idioms according to this definition:

- (10) a. get X's goat
- b. miss the boat
- c. a little bird(ie) told X Y (not a phrasal idiom)

Get X's goat is a phrasal idiom because the verb *get* does not have its literal meaning of acquisition in this idiom, nor does *X's goat* refer to an animal. Just when combined, they produce the meaning 'drive X to anger/annoyance'. The verb *get* does not have this meaning in any other context, nor does *X's goat*. This particular idiomatic meaning only arises when this particular V and this particular NP combine together; it is not there with different choices of NP or different choices of V. So, none of *get X's sheep/cow/hen* or *have/take/bring/steal X's goat* has the same meaning as *get X's goat*.

A good illustration of the role of combination in determining meaning is the three idioms *hit the sack* ('go to bed'), *get/give the sack* ('be fired or fire someone'), and *leave holding the sack* ('abandon to take all the responsibility'). Here, the NP *the sack* has three different meanings triggered by the other words that it combines with. It does not have any of these interpretations without those other words.

Moving on to the second illustrative example, the expression *miss the boat* ('miss an opportunity') is also a phrasal idiom, although the verb, *miss*, has its purely literal meaning in the expression (Riehemann 2001). This idiom only occurs with these particular words; crucially, *the boat* does not occur with the same meaning with other words. For instance, *caught the boat* does not mean that the opportunity was seized. Note that the literature on idioms sometimes distinguishes non-literal expressions like *the boat* from literal expressions like *miss* within idioms (e.g., Nicolas 1995, Riehemann 2001), but this literature also does not make anything of this distinction, and treats both types in the same manner. I will do the same here, although I will attempt to use only those idioms that are entirely non-literal.

In contrast with the first two phrases, the expression *a little bird(ie) told X Y* ('an anonymous source informed X of Y') is not a phrasal idiom. The NP *little bird* can mean an anonymous source of information with verbs other than *tell*. It is attested in *a little bird said/emailed/broadcast/leaked/etc.* The NP even occurs without any verb of communication at all, as in *Had Varys's little birds failed him for once?* (George R.R. Martin, *A Clash of Kings*). This means that only the NP, *little bird(ie)*, has this special meaning, and the verb is not included.

In the rest of this paper I will make use of phrasal idioms identified according to the criteria listed above. I will attempt to use only idioms where all parts are interpreted in a non-literal way, simply to try to make the examples the strongest possible illustration of the possibilities open to idioms. This does not imply that there should be a theoretical distinction between phrasal idioms and other types of fixed expressions; in fact I believe they all work in the same way (see Bruening 2015). I will come back to a theory of idioms (and fixed expressions) and how they work in section 7.

The literature on idioms also sometimes distinguishes idioms according to how compositional they appear to be. Some idioms seem to be decomposable in some way. For instance, in *miss the boat*, the NP *the boat* seems to be a metaphor for an opportunity, while the verb has its regular meaning. In *spill the beans*, *the beans* appears to be a metaphor for a secret or information, while the verb *spill* seems to have a metaphorical use of divulging information (which it may have in other fixed expressions, like *Spill it!* and *spill X's guts*). Nunberg, Sag, and Wasow (1994) call such partially decomposable idioms *idiomatically combining expressions*. In contrast, the parts of other idioms, like *kick the bucket* ('die') and *saw logs* ('sleep snoring loudly'), do not seem to have meanings independent of the whole. Nunberg, Sag, and Wasow (1994) call such idioms *idiomatic phrases*. The distinction is sometimes hypothesized to correlate with how syntactically flexible an idiom is. For instance, the partially decomposable *spill the beans* can be passivized, while the non-decomposable *kick the bucket* cannot.

This distinction will take on some importance in sections 3.3 and 6.2, although it appears to me that there is no clear distinction between these two classes, but more of a continuum. For instance, although *miss the boat* appears to be readily decomposable, it does not passivize easily. Conversely, although the pieces of *the shit hit the fan* are not easy to assign metaphorical meanings to, the NP *the shit* readily undergoes syntactic processes like raising. Similarly, although *kick the bucket* is generally regarded as completely inflexible and non-decomposable, we will

see in section 3 that it can participate in pronominal anaphora. I therefore do not believe that decomposability should be assigned any theoretical significance. (But see sections 3.3 and 6.2 for further discussion.)

With these preliminaries, I turn to syntactic dependencies and show that pieces of idioms can readily participate in many types of dependencies other than movement. The only constraint that they must obey is the CIC. It is not the case that a piece of an idiom must be related to the rest of its idiom by a movement dependency.

3 Anaphora and Ellipsis

I begin with anaphora. This section shows that, so long as the CIC is respected, an idiom chunk can antecede an anaphoric expression. Anaphoric expressions include pronominals, elided VPs (and NPs), and the VP anaphor *do so*. It is not the case that a position anteceded by an idiom chunk can only be the trace of movement of that idiom chunk.

3.1 Pronominal Anaphora

Idiom chunks that are NPs can antecede overt pronominal forms. Particularly productive contexts for this are ones like the following:

- (11) a. I thought I would *miss the boat*, and *miss it* I did.
b. A: *The shit will hit the fan* tonight. B: Yes, **it** will.
c. A: *The cat's out of the bag*. B: **It** certainly is!
d. A: It looks like *the beans* have been *spilled*. B: Yes, **they** have.
e. When *the shit hit the fan*, and **it** did, we had a lot of shoulders to offload responsibility on.
(http://www.shattercolors.com/interviews/author_saknussem.htm)

Often the pronoun is also accompanied by VP ellipsis, discussed further below. This is not necessary, however, as the following attested examples show:

- (12) a. I completely agree he should have *spilled the beans*, but it seems he *spilled them* as a form of pre-emptive justification rather than with any sense of regret (at least based on his interview comments so far).
(<http://www.hkmdb.com/phpbb/viewtopic.php?f=21&t=49199>)
b. I'm pretty sure that when it's my time to *kick the bucket*, they'll let me *kick it* and get out of the way. It's not so bad to exit the way Lorraine did, fast, and probably without pain. Her family says she had it her way.
(<http://carolynrosenblatt.com/blog/>)
c. The 'Green Grape' and 'Cherokee Purple' [varieties of tomato] *bit the dust* and *bit it* HARD but all the other [sic] persevered.
(<http://www.joiecrary.com/blog/>)
d. If you wanna *bury the hatchet*, let's *bury it* right now.
(<http://www.cagematch.net/?id=93&nr=631>)
e. You see, while VPs can *pull strings*, they can't *pull them* in a cost free way.
(<http://evilhrlady.org/2010/01/will-my-high-level-connections-overcome.html>)

Note that even in *kick the bucket*, the prototypical example of an inflexible and non-decomposable idiom (Nunberg, Sag, and Wasow 1994), *the bucket* can antecede a pronoun (12b).

It would not be possible to analyze these examples as involving a movement dependency, from the position of the pronoun to the position of its antecedent. In (12b), for instance, such an analysis would require movement from a main clause (an embedded one) into an adjunct clause adjoined to that clause. No theory of movement that I am aware of would permit such an operation. In (11b), a movement analysis would require movement from one

speaker's utterance into another speaker's utterance. Movement, being part of sentence grammar, cannot do this, in any theory. It must be the case, then, that idiom chunks can participate in the syntactic relation of pronominal anaphora in addition to syntactic movement. This becomes relevant when there are two competing analyses of some phenomenon, one of which involves movement and the other of which involves binding of a (null) pronoun. In such a situation, idioms cannot decide between the two analyses, because they are compatible with both (see below on raising versus control and competing analyses of relative clauses and tough-movement).

In all of the above examples, both the antecedent idiom chunk and the pronoun occur in the context of the rest of the idiom. These examples therefore respect the CIC and are well-formed. If one of the contexts requires a different interpretation, the CIC is violated and the sentence is not well-formed on the idiomatic interpretation:

- (13) a. Bill called plays, then *threw Peterson under the bus*. He should have *thrown* himself *under it*.
 (http://www.tulsaworld.com/blogs/sports/michaelpeters/editor-s-desk-bill-blankenship-gives-article_f6308568-81fc-11e3-8d1b-0019bb30f31a.html)
- b. Bill *threw Peterson under the bus*. #He should have crawled under *it* himself.
- (14) a. I'm pretty sure that when it's my time to *kick the bucket*, they'll let me *kick it* and get out of the way.
- b. # I'm pretty sure that when it's my time to *kick the bucket*, they'll let me have *it* and get out of the way.

This means that idiom chunks are not limited to movement dependencies, but they do obey the CIC. A pronoun can receive an idiomatic interpretation when its antecedent is an idiom chunk, so long as its local context does not require it to be interpreted in a way that is inconsistent with the idiomatic interpretation.

Idiom chunks can also be e-type pronouns, where they have a quantificational interpretation but the quantifier that binds them (in this case an idiom chunk) does not c-command them (Evans 1980):

- (15) a. Every manager who truly *calls the shots* in his office wants everyone to know that he *calls them*.
- b. If you feel secure in your position, you should *pull every string* available to you. If you don't, you shouldn't *pull them*.
- c. Every comedian who was worried about *beating a dead horse* of a joke did *beat it*.
- d. Every mob enforcer who was able to *do a number on* his rival did *do it* on him.

They can also antecede paycheck pronouns (Karttunen 1969):

- (16) a. The worker who *lost his cool* yesterday is better off than the one who *lost it* today.
- b. A POW who *spills his guts* only after hours of torture is more honorable than one who *spills them* immediately.

All of the above examples show that idiom chunks are not limited to movement dependencies. They can also antecede pronouns, with all of the interpretations that are open to pronouns in general. The only requirement is that they obey the CIC.

Before moving on, it should also be noted that idioms can participate in *one* anaphora:

- (17) a. Government schooling *did a real number on* you. Dropping out *did a bigger one* on you, it seems.
 (https://www.reddit.com/r/Bitcoin/comments/35g4nj/the_government_of_denmark_wants_people_to_stop/)
- b. Because just as I *have a bone to pick with* the homosexuals, I also *have one to pick with* my fellow Christians.
 (<http://theredstringblog.com/tag/christianity>)

3.2 VP Ellipsis and VP Pro-Forms

Elided VPs and VP pro-forms, like pronouns, can take antecedents from different sentences and even across speakers (e.g., Hankamer and Sag 1976). This makes it impossible to claim that there is a movement relation between the elided VP or VP pro-form and its antecedent. Movement relations cannot cross sentence boundaries.

Now consider the following:

- (18) a. Every *string* will be *pulled* that can be.
 b. A: *The cat* is *out of the bag*. B: I know it is.
 c. A: *The shit* will *hit the fan* tonight! B: Yes, it will./ It already has.

Many of the examples in the previous subsection also included VP ellipsis of part of an idiom.

VP ellipsis in these examples is not problematic for the view that takes idioms to be a movement diagnostic, if lexical material is actually present but unpronounced in VP ellipsis. B's response in example (18b) would have the following representation, with the strikethrough indicating lack of pronunciation (PF deletion, for instance, as in Merchant 2001):

- (19) I know it is ~~out of the bag~~.

Since the lexical material *out of the bag* is present here, it should be able to contribute to the idiomatic meaning (but something must be said about the pronoun).

VP pro-forms like *do so* behave exactly like VP ellipsis with idioms, and they *are* problematic for the movement view, in exactly the same way as pronouns:¹

- (20) a. A: I'm going to have to *pull* some *strings* to get you what you want. B: Then do so!
 b. He let *the cat out of the bag* on national television before she was able to do so.
 c. He wants to *call the shots* around here but he has never succeeded in doing so.
 d. *The shit hits the fan* with astonishing regularity around here. How can it do so so frequently?

Do so is generally regarded as a pro-form lacking the lexical content and syntactic structure of its antecedent, analogous to a pronoun (e.g., Bresnan 1971, Hankamer and Sag 1976, Fiengo and May 1994, Hallman 2013). In (20a), then, B's sentence is just the following, with a covert imperative subject:

- (21) Then ~~you~~ do so!

Here, there is no constituent *pull strings*, and so there should be no idiomatic meaning. Nevertheless, this is exactly what speakers recover as the meaning of the sentence.

In (20a), the entire idiom is the antecedent of *do so*. In (20d), though, only part of it is: the antecedent of *do so* is just *hit the fan*. The idiom can be spread across two different pro-forms, one a pronoun (*it*) and the other *do so*. Both types of pro-forms are problematic for the view that idioms can be interpreted only on the basis of a movement derivation.

The above examples all respect the CIC. In contrast, if one of the clauses requires an interpretation that is incompatible with the meaning of the idiom, the result is ill-formed:

- (22) # *The shit hits the fan* with astonishing regularity around here and the manure does so too.
 (23) a. *The cat got her tongue* on one talk show on Thursday night, and it did so again on another talk show on Friday night.
 b. # *The cat got her tongue* on one talk show on Thursday night, and the dog did so on another talk show on Friday night.

The facts are exactly as predicted by the CIC. They are not as predicted by the view that the only element that can be anteceded by an idiom chunk is a movement trace.

I note before moving on that idioms can also participate in antecedent contained deletion, a variety of VP ellipsis (Bouton 1970, Sag 1976):

- (24) a. She can *pull* every *string* that we can.
 b. We can *make* all the *headway* they can.

¹These data directly contradict Hallman (2013), who claims that idioms are not well-formed with *do so*.

Idioms can also participate in NP ellipsis:

- (25) a. I should've *cut* him *some slack*, ... or *cut* myself *some*.
(<http://www.goldwingfacts.com/forums/162-altered-induction/358561-g11000-single-carb-conver.html>)
- b. Other cars in this segment *cut* some *corners*. We didn't *cut* any: that's why we included cruise control and fog lamps with corner lights.
(<http://www.seat.com.cy/content/cy/brand/en/models/alhambra/versions/style.php>)

Once again, idioms are able to participate in the full range of anaphoric possibilities, even ones crossing sentence boundaries. They are not limited to movement derivations.

3.3 Apparently Problematic Cases

Attested examples like the following might initially seem to be problematic for the CIC:

- (26) a. They had the *strings* and *pulled them*. (Riehemann 2001, 75, (73e))
- b. Back in her days of the studio system era, the tycoons held all the *strings*. And *pulled them* so hard, they could yank Bette Davis or Joan Crawford out of films and punish Clark Gable for some sin against them by sending him into some no-hope Poverty Row film.
(<http://www.crawleyscastingcalls.com/index.php/component/intro/?Itemid=75>)
- (27) a. **HERE'S THE HATCHET, DO YOU BURY IT?**
(<http://strippedbar3.blogspot.com/2004/07/heres-hatchet-do-you-bury-it.html>)
- b. This doesn't mean that people have to ultimately agree or even like each other, but they have to agree on what the *hatchet* is and *bury it* completely.
(<http://innovergent.com/2014/12/12/sons-anarchy-gave-tips-manage-conflict-consulting/>)
- (28) a. Well, if you have the *beans*, then *spill them*.
(<http://www.rcrawler.com/forum/tires-wheels/196561-new-hb.html>)
- b. **The beans. Spill them.**
(<http://suptg.thisisnotatrueending.com/archive/31010161/>)

In these examples, part of an idiom occurs as a thematic argument in a clause that does not include the rest of the idiom. A pronoun that takes that idiom chunk as its antecedent does occur with the rest of the idiom.

The pronoun is not problematic, it is just another instance of the ability of idiom chunks to participate in pronominal anaphora (see above). The problem is the antecedent of the pronoun. It is an idiom chunk that occurs as an argument in a clause that does not include the rest of its idiom, and therefore might be expected to violate the CIC.

In fact, though, these cases are predicted by the CIC, and provide important evidence in favor of the theory presented here. The CIC only requires that an idiom chunk not be assigned an interpretation that is inconsistent with its use in the idiom. It does not require that the idiom chunk be assigned *no* interpretation, or occur in a non-thematic position. This is a crucial difference between the CIC and the movement view. In the theory here, all that matters is that the interpretation that is being assigned is consistent with the idiom. In contrast, the idea behind the movement view is that an idiom chunk can only move to a position that is assigned *no* interpretation at all, so that it is only interpreted with respect to its idiom.

These cases are therefore devastating to the movement view. The idiom chunk appears in an argument position in a completely different sentence from the rest of the idiom. It *must* be occupying a thematic position that is being assigned an interpretation. In contrast, all of these examples can be shown to be compatible with the CIC. These are the types of idioms that speakers seem to be able to decompose in such a way that they can assign meanings to the component parts. Nunberg, Sag, and Wasow (1994) discuss this with *spill the beans*, suggesting that speakers interpret *spill* as 'divulge' and *the beans* as 'the secret' or 'insider information'. In *pull strings*, *strings* is interpreted as 'connections' or 'influence', and *pull* as 'use/manipulate' (the metaphor here is that of a puppeteer). In *bury the*

hatchet, *the hatchet* is a metaphor for a grudge or conflict, and *bury* is ‘lay to rest, end’. In the two examples of *pull the strings* in (26), the NP occurs with the verbs *have* and *hold*. These do not assign any meaning to their direct object that is inconsistent with the meaning ‘connections, influence’: one can have and hold onto connections and influence. These two examples are therefore consistent with the CIC. In *here’s the hatchet* (27a), there is no meaning that is inconsistent with the meaning of a grudge or conflict. Similarly for *agree on what the hatchet is* (27b), which clearly refers to the conflict. The same holds for the two examples of *spill the beans* in (28), where the interpretation is consistent with ‘secret/information’.

These cases therefore actually provide support *for* the CIC, and they are problematic for many theories of idioms (for an overview of some of these theories, see section 7). In section 7.4, I show how the theory of idioms proposed here explains such cases, and predicts correctly that all parts of the idiom must appear in the linguistic context (so, *Here’s the hatchet* would not have the idiomatic meaning by itself, in the absence of *bury*). In section 4, next, I show how some similar control cases also provide crucial support for the CIC, and contradict the received view that idioms may not participate in control.

4 Control and Raising

I return now to control versus raising. In analyses of control that posit a null pronominal (“PRO”) as the subject of a non-finite clause, control is a variety of pronominal anaphora, which we have now seen that idioms can participate in. We might therefore expect them to be well-formed in control relations.

As stated in the introduction, however, numerous publications have presented contrasts like the following to show that idioms distinguish control from raising:

- (29) (Davies and Dubinsky 2004, 8, (23a–b))
- a. *The cat* seemed to be *out of the bag*.
 - b. *The cat* tried to be *out of the bag*. (*idiomatic)
- (30) (Postal 2004, 93, 26b–c)
- a. They believe *the shit* to have *hit the fan* yesterday.
 - b. * They convinced *the shit* to *hit the fan* yesterday.

This claim is usually based on two assumptions: first, that idiom chunks cannot receive any sort of interpretation (or thematic role) other than the interpretation they receive from the rest of their idiom; and second, that the idiomatic interpretation requires that at some level of representation, the idiom chunk is in a sufficiently local relationship with the rest of its idiom so that it can receive that interpretation. Jointly, these assumptions rule out an idiom chunk as a thematic argument of a control verb, controlling a null pronominal (PRO) in the non-finite clause. If it were the thematic argument of a control verb, it would receive an interpretation independent of and distinct from its idiomatic interpretation (from the control verb that it is a thematic argument of); and it would not be able to receive the idiomatic interpretation through its control link with PRO, because only movement dependencies can put the idiom chunk in the required local relationship with the rest of the idiom.

As also stated in the introduction, however, it is possible to find examples of idiom chunks participating in control. I have found two subject control verbs—*want* and *wait*—and one object control verb—*force*—with multiple attested examples of idioms:

- (31) a. To some, the lesson is that no matter how happy the tune may be, eventually, *the piper* wants to be *paid*. (Riehemann 2001, 62, (60a))
- b. *Corners* always want to be *cut*, people naturally want to do what is easy, and without strong leadership to remind people. . .
(http://gembapantarei.com/2010/09/how_to_sustain_a_lean_culture_after_10_years.html)
- c. If the situation’s so bad that you have to rock the boat, then, unless *the boat* wants to be *rocked*, it may be time to move on.
(<http://www.hbcusports.com/forums/threads/principles-over-money.101269/page-2>)

- (32) a. Eat, drink and be merry, Barry and Mechille, for the party is almost over and *the piper* waits to be *paid*.
(<https://motuspnyxe.wordpress.com/2010/08/28/our-recovery-summer-we-could-use-those-jaws-of>)
- b. *The plug* is waiting to be *pulled* once all the 5600 stock is gone. . .
(<http://www.oesf.org/forum/lofiversion/index.php/t1554.html>)
- c. Why? . . . cus as I said, *corners* are waiting to be *cut* where there is money to be made.
(<https://www.caravanbanter.co.uk/9468-re-undertaking-motorway-35.html>)
- d. And then there is the garam gossip ‘n gupshup of the girls’ Freaky Friday night out at Shoba’s. *The beans* are waiting to be *spilled* [smile emoji].
(<https://desiwisecracks.wordpress.com/random-rants-2/>)
- (33) a. In Montana, the temptation to buy popularity in the short term has bought nothing in the long term because routine obligations for services have forced *the piper* to be *paid*.
(http://billingsgazette.com/news/opinion/guest/guest-opinion-cycle-of-surplus-and-shortfall/article_604bcc09-0bd4-5be2-a49f-1eda79f1a286.html)
- b. A spokesman for the manufacturer said lack of testing on the circuit prior to the event forced *the plug* to be *pulled* at the eleventh hour.
(<http://cars.aol.co.uk/2011/07/04/nissan-forced-to-pull-reversing-stunt-from-fesitval-of-sp>)
- c. . . but budget constraints forced *corners* to be *cut* elsewhere, resulting in a 5atm watch.
(<http://www.christopherwardforum.com/viewtopic.php?f=1&t=33473&start=30>)

Such examples are problematic for the longstanding view in the literature on raising and control. Apparently, an idiom chunk *can* appear as a thematic argument of a control verb, controlling a null pronominal in an infinitival complement.

These verbs do not appear to have been miscategorized as control verbs. They are still incompatible with expletive *there*, another diagnostic of control as opposed to raising:²

- (34) a. * There wants to be a party.
b. * There waits to be an inquiry.
c. * John forced there to be more exams taken in his class than in any other. (Akmajian and Heny 1975, 329)

They are also incompatible with extraposition *it*:

- (35) a. * It wants to be true that diamonds are not actually rare.
b. * It waits to be recognized (by everyone) that diamonds are not actually rare.
c. * The cartel forced it to be true that diamonds are not actually rare.

These verbs do appear to be control verbs (as argued for *force* by Akmajian and Heny 1975, 329 and much other literature). What distinguishes them from *try* and *convince* is that they do not necessarily require any particular interpretation of their subject (in the case of *want* and *wait*) or object (in the case of *force*). As was explained in the introduction, one can force a state of affairs without acting directly on the direct object of *force*, for instance. This distinguishes *force* from *convince*:

- (36) a. Those changes forced the bill to be sent back to the House one more time.
b. * Those changes convinced the bill to be sent back to the House one more time.

Convince requires an object that is sentient and capable of being swayed by arguments. In contrast, *force* does not make any requirements of its direct object.

We can see a similar distinction between *want/wait* and *try*:

²No attested examples of this sort appear on internet searches. In fact, all of the hits for “forced there to be” are examples of ungrammatical sentences from the linguistics literature.

- (37) a. I applied a coat of Ultra Grip, as sometimes metal doesn't want to be painted.
 b. The metal was waiting to be painted.
 c. * Sometimes metal tries not to be painted.

As was explained in the introduction, *want* seems to be able to refer to the perceived state of affairs in the world, without necessarily attributing desires to its syntactic subject. As for *wait*, its subject can be completely inert.

Given these semantic facts, idiom chunks can appear with control verbs like *want*, *wait*, and *force* without violating the CIC, even though they are thematic arguments of these verbs. They cannot appear with *try* or *convince* or similar control verbs, or they do violate the CIC, as in (29b) and 30b). Importantly, it is not the thematic status of the position of the idiom chunk that matters: an idiom chunk can be a thematic argument of some predicate that is not the rest of its idiom, so long as that predicate does not assign it an interpretation that is incompatible with the idiomatic interpretation it receives through an anaphoric linkage to the rest of its idiom (here, control). We saw the same thing with cross-sentential anaphora in section 3.3.

Furthermore, the ability of idiom chunks to appear with some control verbs, in particular the class represented by *want*, *wait*, and *force*, is incompatible with the view that idiom chunks can only be related to the rest of their idiom by movement. With these verbs, they are related to the rest of the idiom by control, not movement.

The ability of idiom chunks to participate in control can be further exemplified by looking at control into adjunct clauses. The null subject of a non-finite adjunct clause is controlled by the subject of the clause it is adjoined to. The verb of that clause can be anything, even an idiomatic verb. The CIC then predicts that we should be able to find idiom chunks controlling into such clauses, and indeed they can (I represent the null subject as PRO):

- (38) a. *The shit will hit the fan* in our office without PRO *hitting* it in yours.
 b. Could *the cat* have been let *out of the bag* without PRO being let *out* deliberately?
 c. Could *the beans* have been *spilled* without PRO being *spilled* deliberately?
 d. *Headway* was *made* without PRO seeming to be *made*.
 e. *Strings* can be *pulled* without PRO being *pulled* in any obvious way.

Note that in these examples the idiom chunk is interpreted purely idiomatically in both the matrix clause and the adjunct clause. This satisfies the CIC, as there is no interpretation required of the idiom chunk that is not compatible with its meaning in the idiom. If such an interpretation is required in either clause, the sentence is not acceptable, as was pointed out by an anonymous reviewer:³

- (39) a. # *The shit* will be a problem even without PRO *hitting the fan*.
 b. # *The shit* will *hit the fan* without PRO ever being a problem.

As all of these examples show, idiom chunks can control the subject of a non-finite clause, so long as they respect the CIC. This contradicts the view dating to at least Postal (1974), according to which idiom chunks cannot participate in control.

Note that idioms differ from the expletive *there*, which cannot control the null subject of a non-finite adjunct clause:

- (i) There will be a problem.

The examples in (39) ought to then be a problem for the CIC, since *the shit* does not have to be assigned any interpretation that is incompatible with the idiom *the shit hit the fan*. Note, however, that there is a difference between (i) and any other NP as subject of *will be a problem*: (i) asserts the existence of a problem, but any other subject has *problem* predicated of it. The existential interpretation is particular to the expletive construction. In (39), then, *the shit* is necessarily being asserted to be a problem, and this is incompatible with its use in the idiom, where it is non-referential and cannot be a problem on its own (the fact that it hits the fan can be a problem, but *the shit* cannot itself be a problem).

- (40) a. * There occurred three more accidents without PRO being any medical help available on the premises.
(Haegeman 1994, 279, (46g))
- b. * Next to the cauldron there appeared a witch without PRO appearing a black cat, too.

The expletive *there* also cannot antecede pronouns (different occurrences of *there* are independent of each other). We saw above that it also cannot be an argument of a control verb. This expletive *can* undergo movement (A-movement, at least). It therefore does appear to be a valid movement diagnostic, as opposed to idiom chunks (see more on this below).

Note furthermore that there is a movement analysis of control (Hornstein 1999, 2001), according to which the null subject of an infinitive is actually the trace of its controller. This analysis would have to be supplemented with something like the CIC, in order to differentiate *try* and *convince* from *want*, *wait*, and *force*, and to differentiate (38) from (39). In other words, the CIC does all the work with idioms, regardless of how we analyze control. Limiting idiom chunks to movement either gets the facts wrong, or does no work. Note also that proponents of the movement theory of control could not point to the acceptability of some idioms in control as an argument for the movement theory of control. We saw in the preceding section that idiom chunks can also antecede pronouns. Idioms are therefore also compatible with the PRO theory of control, as well as the movement theory of control. Both just need to be supplemented with the CIC.

5 Parasitic Gaps

Parasitic gaps as in (41) further illustrate the importance of the CIC, and they will also make an important point for more contentious cases in section 6.

- (41) How many reports did she file *t* without reading *pg*?

In parasitic gap constructions, there are two gaps related to a single filler. One, indicated above with *t*, is a gap that is legitimate on its own. The other, *pg* above, occurs in a position that is not a legitimate location for such a gap (inside a movement island) without the other gap. In the example above, the parasitic gap is inside an adjunct clause. Adjunct clauses are often islands to movement. The gap inside this island is parasitic on the existence of the legitimate gap outside the island.

There are two main analyses of parasitic gaps, the null operator analysis and the sideward movement analysis. In the null operator analysis, the real gap outside the island is the trace of movement, while the parasitic gap inside the island is the trace of a null operator that moves within the island (e.g., Chomsky 1986, Nissenbaum 1998). Example (41) would be analyzed as the following, with a null operator moving within the adjunct clause:

- (42) [How many reports]₁ did she file *t*₁ [OP₂ without reading *t*₂]?

There is no island violation here because the null operator does not cross the island boundary. The two movement chains need to be unified in some way; different proposals have been made for doing this (e.g., Chomsky 1986, Nissenbaum 1998).

The sideward movement analysis, in contrast, posits a movement relation between the wh-phrase and both the real gap and the parasitic gap (Nunes 2004). The wh-phrase starts out in the position of the parasitic gap, and then moves “sideward” into a syntactic structure that is not yet connected to the one it started out in. It moves to the position of the real gap, object of *file* in this example, in the main clause being constructed in a separate workspace. From there it undergoes movement to its surface A-bar position. Once the two trees are connected (by the adjunct clause adjoining to the main clause), the wh-phrase ends up c-commanding all of its traces, satisfying constraints on movement. Movement never crossed an island boundary, either, because the sideward movement sidestepped the island boundary.

With these two analyses in mind, let us look at idiom chunks and parasitic gaps. Chunks of idioms are known to be able to undergo A-bar movement like wh-movement and relativization (see section 6.2). In doing so, they can also license parasitic gaps:

- (43) a. We were amazed at the *strings* she managed to *pull t* without seeming to *pull pg*.
 b. These are the kind of *strings* that people who *pull pg* usually end up regretting having *pulled t*.
- (44) a. How much *headway* can we *make t* without appearing to *make pg*?
 b. That's the kind of *headway* that people who *make pg* always want to brag about having *made t*.

Notice that in these examples, both the real gap and the parasitic gap occur with the rest of the idiom. The idiom occurs twice. This satisfies the CIC, because there is no position where the idiom chunk is being assigned an interpretation that is inconsistent with the interpretation it requires for the idiom. If either the parasitic gap or the real gap *is* assigned some other interpretation, the result is ill-formed:

- (45) a. # That's the kind of *headway* that people who bring about *pg* always want to brag about having *made t*.
 b. # How much *headway* can we *make t* without appearing to be excited about *pg*?
- (46) a. # That's the kind of *headway* that people who *make pg* always want to brag about having accomplished *t*.
 b. # How much *headway* can we brag about *t* without actually *making pg*?

This contrast is exactly as expected by the CIC. It is not expected by the view that movement is the only possible syntactic dependency for idiom chunks, in either the null operator or the sideward movement analysis. In the null operator analysis, the parasitic gap is not related to the overt idiom chunk by movement. It is the trace of a null operator, not the trace of the idiom chunk. This analysis therefore expects parasitic gaps to always be ill-formed with idiom chunks; the fact that they are acceptable in (43–44) is unexpected. Now, we could use whatever mechanism unifies the operator chain with the main filler-gap chain to explain why the idioms are acceptable in (43–44); the two chains are unified and become a single chain. However, once we do that, we need something like the CIC to rule out (45–46). And once we add the CIC, there is no need for a constraint limiting idiom chunks to movement relations. The CIC does all the work.

In the sideward movement analysis, the idiom chunk is related by movement to both the parasitic gap and the real gap. If idiom chunks can retain their idiomatic interpretations just in movement relations, they ought to be able to receive their idiomatic interpretation in either the real gap or the parasitic gap position. The examples in (45–46) are then unexpected. In order to rule these out, the sideward movement analysis would have to add a constraint like the CIC in addition. Once we do that, however, we do not need the constraint that idiom chunks can only be related to the rest of their idiom by movement. The CIC again does all the work.

What parasitic gaps show us is that we need the CIC, regardless of how we analyze parasitic gaps. If we have the CIC, there is no need to limit idiom interpretations to movement dependencies, and doing so leads to the wrong results in the null operator analysis.⁴

Importantly for the next section, if we adopt the null operator analysis, then idiom chunks can be related to the rest of their idiom through the mediation of a null operator. In the preceding section, we saw that they could also be mediated through the null pronominal PRO. Putting these together, we have a situation like this:

- (47) [_{XP} NP₁ X ... [_{YP} *ec*₁ Y]] (linear order irrelevant)

In the above, Y is some predicate that is part of an idiom. NP₁ can be the rest of idiom that goes with Y so long as X does not assign an interpretation to NP₁ that is incompatible with the idiom. The empty category *ec*₁ can be

⁴In some analyses of across-the-board (ATB) movement in coordination, the second (and subsequent) gaps are not directly related to the overt antecedent (e.g., Munn 1993, Zhang 2010). Idioms can also participate in ATB movement:

- (i) a. They were surprised at the *headway* that we *made* but they didn't *make*.
 b. We were surprised at the *strings* that he consented to *pull* but she refused to *pull*.

In the analyses of ATB movement just cited, these examples would also be incompatible with the assumptions behind using idioms as a movement diagnostic.

a trace, a null operator, or PRO. In the next section, we look at two cases of exactly this situation, where idioms have been used as an argument for a movement analysis over a different sort of empty category analysis. What we have seen here renders this argument void: idioms are compatible with the empty category being a null operator or a null pronoun in addition to a trace.

6 Contentious Cases: Tough-Movement and Relative Clauses

Both tough-movement and relative clauses have been analyzed as either direct movement, or movement of a (null) operator. As we will see here, idioms cannot be used to decide between these two analyses.

6.1 Tough-Movement

Despite some claims to the contrary (e.g., Chomsky 1981), it is generally agreed that pieces of idioms can undergo tough movement. Here are some examples from the literature:

- (48) (Berman 1973, (4–5))
- a. *Headway* should be easy to *make* in cases like this.
 - b. *Allowances* are easy to *make* for the very young.
 - c. *The hatchet* is hard to *bury* after long years of war.
 - d. *The ice* was hard to *break* at first.
- (49) (Ruwet 1991, 186)
- a. *The curtain* will be difficult to *lower on* this scandal.
 - b. *A can of worms* is often easier to *open* than one expects.
 - c. ... *the line* can be hard to *toe* when push comes to shove.
- (50) (Nunberg, Sag, and Wasow 1994, (49))
- a. Some *strings* are harder to *pull* than others.
 - b. *This boat* is very easy to *rock*.
 - c. *That line* is hard to *swallow*.
 - d. *This barrel* is, unfortunately, very easy to *scrape the bottom of*.
 - e. *His closets* would be easy to find *skeletons in*.

Importantly, the idiom chunk cannot occur as the subject of the tough-adjective without the rest of the idiom:

- (51) a. *Headway* should be easy to *make* in cases like this.
b. **Headway* should be easy.
- (52) a. *The hatchet* is hard to *bury* after long years of war.
b. *The hatchet* is hard. (*idiomatic)

(This is predicted by the account of idioms proposed in section 7.)

It should be easy to see how the CIC permits idiom chunks in tough-movement, given that tough-movement is generally regarded as synonymous with its expletive counterpart. For instance, (53a) and (53b) are semantically equivalent:

- (53) a. *Headway* should be easy to *make* in cases like this.
b. It should be easy to *make headway* in cases like this.

Since the idiom chunk as subject of the tough-predicate does not appear to be assigned any interpretation at all, it is not assigned any interpretation that is incompatible with the interpretation it requires for the idiom.

There are two main analyses of tough-movement. One posits base-generation of the subject of the tough-predicate, with movement of a null operator within the non-finite clause. The subject of the tough-predicate is related to the gap in the non-finite clause not through movement, but through predication via the null operator (Chomsky 1977, 1981).

Opposed to the null operator analysis are different versions of a direct movement approach to tough-movement. These have the subject of the tough-predicate moving directly from the gap within the non-finite clause to the subject position, typically by way of an A-bar position at the edge of the non-finite clause (Brody 1993; Hornstein 2001; Hicks 2003, 2009; Hartman 2009, 2012; Pesetsky 2012). This second type of analysis has to reject or reformulate the generally assumed ban on *improper movement*, which rules out A-bar movement followed by A-movement, or hypothesize some way to get around it (Hicks 2003, 2009).

The fact that idiom chunks are acceptable in tough-movement might be taken to be an argument for the direct movement approach and against the null operator approach, given the common view that idiom chunks can only be related to the rest of their idiom by a movement relation. In fact, Hicks (2009, 554) presents data from idioms as an argument for a direct movement approach.

However, given the findings of this paper, idioms cannot be used as an argument for a direct movement approach and against a null operator approach. They are compatible with either. In particular, they are compatible with the null operator analysis. This analysis would analyze an idiom example as follows, with movement of a null operator within the non-finite clause:

- (54) Headway should be easy [_{CP} OP₁ PRO to make *t*₁] in cases like this.

The null operator is then linked to the subject of the tough-predicate by predication. Given that there is such a link, the idiomatic reading of *headway* in the context of *make* can be established (see section 7 for details), the same way that an idiomatic reading can be established in cases of control and parasitic gaps, above. The only requirement is that *headway* not be assigned some other interpretation in its own position that is incompatible with the idiomatic interpretation. In tough-movement, the surface subject position seems not to be assigned any interpretation at all, so this constraint (the CIC) is met.

Idioms are compatible with the null operator analysis, then, and they cannot be used as an argument against that analysis and for the direct movement analysis. As we have seen throughout this paper, idioms do not diagnose movement.

While it is not the purpose of this paper to argue for one analysis of tough-movement over another, it is worth pointing out that expletives, mentioned above as contrasting with idioms in being a genuine movement diagnostic, fail to undergo tough-movement:

- (55) a. It's impossible (for you) to stop it from snowing in the Himalayas.
b. * It is impossible (for you) to stop from snowing in the Himalayas.
- (56) a. The police believe there to have been a crime committed.
b. * There is hard to believe to have been a crime committed. (Chomsky 1981)
- (57) a. It would be impossible for him to cause there to be a car crash.
b. * There would be impossible (for him) to cause to be a car crash.
(cf. Anger and other negative feelings are easy to cause to be displaced.)
- (58) a. It's hard for me not to love it when you are passed over for promotion so often.
b. * It is hard for me not to love when you are passed over for promotion so often.

The evidence therefore seems to tilt in favor of the null operator analysis, an analysis that idioms are perfectly compatible with. (On the general lack of reconstruction in tough-movement, see Postal 1971, 1974; Hicks 2009; Hartman 2012; the only apparent examples of reconstruction, like those in Sportiche 2002 (cited in Hicks 2009) and Pesetsky 2012, all involve picture-NPs, and are probably not reconstruction at all.)

6.2 Relative Clauses

This paper began with some quotations regarding relative clauses, and this section returns to what idioms have to say about their analysis. Contrary to the view espoused in those quotations, they do not argue for one analysis over another.

I repeat one of the quotes from the literature on relative clauses: “If the idiomatic reading of *headway* is available only if *headway* appears inside the complement of *make*, [(1)] requires the raising analysis” (Hulsey and Sauerland 2006, 113).

(59) (= 1) The *headway* that John *made* was impressive.

Let us critically examine this reasoning, beginning by comparing it with the reasoning that the literature on raising and control used to explain why idioms appeared not to participate in control. To repeat, idiom chunks were thought to be bad in control for two reasons: First, idiom chunks cannot receive any sort of interpretation (or thematic role) other than the interpretation they receive from the rest of their idiom. Being the thematic argument of a control verb would violate this constraint. Second, the idiomatic interpretation requires that at some level of representation, the idiom chunk is in a sufficiently local relationship with the rest of its idiom so that it can receive that interpretation. If an idiom chunk were to occur as a thematic argument of a control verb, it would be unable to get into the requisite local relationship with the rest of its idiom, since that can only be accomplished by way of a movement chain.

Note that the literature on relative clauses only adopts the second part of this reasoning, and ignores the first. With relative clauses, the head noun is always a thematic argument of some other predicate. In (1), *headway* is a thematic argument of *was impressive*. This by itself is ill-formed:

(60) * *The headway* was impressive.

The literature on control versus raising concluded that idiom chunks could only undergo movement into a non-thematic position. The literature on relative clauses adopts the view that idiom chunks can only be displaced from the rest of their idiom by movement, but then it runs into a problem by positing movement into a thematic position. It is not at all clear how exactly the head raising account can permit (1) when (60) is not well-formed.

The argument from idioms for the head raising analysis of relative clauses seems to be intended as an argument against an alternative analysis, one very similar to the null operator analysis of tough-movement. This is the relative operator analysis. In the relative operator analysis, what moves is not the head noun itself, but a relative operator (e.g., Chomsky 1977). This approach would treat (1) as follows:

- (61) a. The *headway* that John *made* was impressive.
b. [NP The headway [CP OP₁ that John made *t*₁]] was impressive.

The CP here is simply a predicate, combined by conjunction with the head N, which is also a predicate (resulting in “ $\lambda x.x$ is headway and John made *x*”). The determiner then combines with this nominal predicate. The entire NP headed by the head of the relative clause is base-generated as the thematic argument of the higher predicate.

Given what we have seen in this paper, there is no difficulty for this operator analysis in getting the idiomatic interpretation of *headway* (and *make*). As we have seen, idiom chunks can trigger idiomatic readings in combination with the rest of their idiom when that combination is mediated by a null operator or a null pronominal (or even an overt one). Here, *headway* is interpreted as the object of *make*, exactly like what happens with a head noun that is interpreted purely literally:

(62) The headband that John made was impressive.

Here, native speakers understand that what was made was a headband. That is, the lexical content of the head noun has to be recovered in the gap position (see section 7). If the null operator analysis succeeds at capturing this incontrovertible fact with literal phrases, as most assume that it does, then it succeeds at capturing idiomatic phrases, too. There is no argument from idioms for movement of the head noun itself. Idioms are fully compatible with a null operator analysis.

Above we noted that the head raising analysis runs into the problem of how to interpret the idiomatic head noun in the matrix clause. The null operator analysis runs into the same problem. Whichever theory we adopt, it appears that the view espoused here is in trouble, because in such cases the CIC appears to be violated. Before turning to this problem, though, I would like to reiterate that both analyses succeed and fail equally. They both succeed in constructing the idiomatic reading, but they both fail at explaining how the idiom chunk can be an argument of the matrix predicate.

6.2.1 The Problem of the Matrix Clause

I turn now to showing how idiom chunks with relative clauses are compatible with the CIC. To illustrate the problem in a way that will be useful as we proceed, relative clauses can generally be paraphrased as a conjunction of two clauses, as follows:

- (63) a. The headband that John made was impressive.
b. \approx John made a headband and that headband was impressive.

In our idiom case, this is ill-formed, because one clause requires a non-idiomatic interpretation:⁵

- (64) a. The *headway* that John *made* was impressive.
b. John *made headway* and *that *headway* was impressive.

In the rest of this section, I attempt to defend the CIC in the face of such data. I have found four different sorts of examples involving idiom chunks being shared across two clauses. Each of these four requires a different analysis.

6.2.2 Class 1: Amount

The first class of examples all seem to involve the idiom *make headway*. Here is a selection of such examples:

- (65) a. The headway that we made was satisfactory. (Schachter 1973, 31, (35c))
b. The headway John made proved insufficient. (Sauerland 1998, 63, (53a))
c. The headway on her project Mary had made pleased the boss. (Sauerland 1998, 71, (64b))
d. John was satisfied by the amount of headway that Mary made. (Hulsey and Sauerland 2006, 113, (3a))
e. Mary praised the headway that John made. (Hulsey and Sauerland 2006, 114, (8a))

In all of these cases, what seems to be at issue in the matrix clause is the *amount* of headway, as is explicitly indicated in (65d). I suggest that in these cases, *headway* is not really an argument in the main clause, some abstract noun like *amount* is. (See Sudo 2015 for an analysis of Japanese clausal comparatives that posits a similar silent degree head.) Paraphrases like the following seem to accurately capture what the above examples mean:

- (66) a. We made some amount of headway and that amount was satisfactory.
b. John made some amount of headway and that amount proved insufficient.
c. Mary had made some amount of headway on her project and that amount pleased the boss.
d. Mary made some amount of headway and John was satisfied by that amount.
e. John made some amount of headway and Mary praised that amount.

⁵Depending on how decomposable this idiom is for a given speaker, that speaker may find (64b) more or less acceptable, like the examples in section 3.3. This is predicted by the account in section 7.

On this analysis, the idiomatic noun is interpreted solely within the relative clause. The noun of the matrix clause is not the idiom chunk at all, it is the silent (or overt) N *amount*.

This first class of examples then satisfies the CIC because the idiom chunk is assigned only one interpretation, namely, the idiomatic one. Note that this explanation is independent of the particular analysis of relative clauses: it can be stated in either head-raising or relative-operator terms. On the relative operator analysis, an example like (65a) would be analyzed as follows:

- (67) a. The headway that we made was satisfactory. (Schachter 1973, 31, (35c))
b. The ~~amount of~~ headway [_{CP} OP₁ that we made *t*₁] was satisfactory.

This requires an NP along the lines of, “the amount of *x* such that *x* is headway and we made *x*.” *Amount of x* does not impose any restrictions on *x* except that it be measurable. This requirement is not incompatible with the idiom, since *headway* can be measured in the idiom:

- (68) a. We made a lot of *headway*.
b. We made very little *headway*.
c. We made more *headway* than they did.

Thus, there is no violation of the CIC under this analysis.

6.2.3 Class 2: Propositions

The second class of relative clauses with idiom chunks as heads consists of examples like the following:

- (69) a. They were surprised at the *corners* that we were willing to *cut*.
b. They were surprised at the *strings* that we managed to *pull*.

These do not seem to be appropriately paraphrased with a noun like *amount* or *number*:

- (70) a. They were surprised at the *corners* that we were willing to *cut*. \neq We were willing to cut some number of corners and they were surprised by that number.
b. They were surprised at the *strings* that we managed to *pull*.
 \neq We managed to pull some number of strings and they were surprised by that number.

Instead, appropriate paraphrases of these sentences can be constructed that do not involve splitting the idiomatic phrases:

- (71) a. They were surprised [that we were willing to *cut* such *corners*].
b. They were surprised [that we managed to *pull* such *strings*].

That is, what they were surprised at is not just the noun, but the entire proposition denoted by the noun plus accompanying clause. It appears that the two together are interpreted as a proposition. A similar propositional interpretation arises with non-idiomatic phrases in relative clauses, too:

- (72) a. They were surprised at the offers I turned down.
b. \approx They were surprised [that I turned down such offers].
c. \neq I turned down some offers and they were surprised by those offers.
(73) a. They were surprised at the hypotheses that I rejected.
b. \approx They were surprised [that I rejected such hypotheses].
c. \neq I rejected some hypotheses and they were surprised by those hypotheses.

In these examples, what they were surprised at is not just the noun; it is an entire proposition, denoted by the noun plus its modifying relative clause. The relative clause does not appear to be performing its usual restrictive function here; in (72), they are not surprised at some particular offers, where the particular offers are the ones picked out by the relative clause. What is surprising is the fact that I turned down the offers.

It therefore appears that at least some instances of head nouns plus relative clauses can be interpreted as full propositions rather than as restrictions on NPs. This seems to be what is going on in many instances of idiomatic phrases as heads of relative clauses. The following is a sample of examples from the literature that seem to have this interpretation:

- (74) a. The careful *track* that she's *keeping of* her expenses pleases me. (Schachter 1973, 32, (36c))
b. I was offended by the *lip service* that was *paid to* civil liberties at the trial. (Schachter 1973, 32, (37c))
c. I was shocked by the *advantage* that she *took of* her mother. (Hulsey and Sauerland 2006, (8b)).
d. . . Robert McNamara's new book justified all the *strings* Clinton *pulled* as a young man. . . (Riehemann 2001, 29, (25))
e. Not even all the *strings* that Sounders coach Alan Hinton *pulled* to get Baggio to make an appearance was enough to fill the stadium. (Riehemann 2001, 75, (73d))

All of the above examples are compatible with the clausal interpretation, as the reader can verify by replacing the NP + relative clause with either *the fact that*. . . or a non-finite clause (e.g., *Robert McNamara's new book justified [Clinton pulling all those strings as a young man]*).

If this is correct, then the second class of idiomatic examples is not problematic, either: the idiomatic NP is really only interpreted in the lower clause, with the rest of the idiom, while the argument in the higher clause is the lower clause, and not just the idiomatic head N. The CIC is not violated.

Once again, this analysis can be stated in either the head raising account or the relative operator account. The relative operator account will need some way to transform the whole NP into a proposition, and that mechanism cannot be total reconstruction of the head noun, since that never moved in the relative operator analysis. The head raising analysis could simply apply total reconstruction to the head N, but even that would not be sufficient, as the determiner and possibly adjectival modifiers would be left behind. The head raising analysis would still need some way to get a determiner plus CP to be interpreted as the appropriate proposition. Thus, both analyses would need additional assumptions to make this explanation work.

6.2.4 Class 3: Double Idiom

The third type of example has the idiom in both the main clause and the relative clause:

- (75) a. John *paid* the same *heed* last year that Mary *paid*. (Henderson 2007, 215, (28a))
b. We can *make any headway* you can *make*.
c. James won't *pull any strings* that Jill is not willing to *pull*.

This is expected to be possible, because the idiom chunk has its idiomatic interpretation in both clauses. The CIC is respected, since no interpretation is being assigned that is inconsistent with the idiom.

6.2.5 Class 4: Decomposable Idioms

The fourth type of example is similar to the examples from section 3.3, where an idiom chunk truly does occur in a thematic argument position that is not part of the idiom. I repeat one of those examples:

- (76) They had the *strings* and *pulled* them. (Riehemann 2001, 75, (73e))

As described in section 3.3, this is possible with decomposable idioms, where the parts can be assigned metaphorical meanings. Here, *strings* is a metaphor for connections or influence, and that is something that one can have. *Strings* can therefore appear as the thematic argument of the verb *have* without violating the CIC.

Attested examples of infinitival relative clauses with idiomatic heads seem to be part of the same pattern:

- (77) a. Why would a reputed master politician fail to find one *string* to *pull*? (Riehemann 2001, 75, (73a))
 b. But Lee has even more *strings* to *pull* ... (Riehemann 2001, 76, (73g))

In these examples, the head noun *strings* is an argument of *find* and *have*, both of which are compatible with the idiomatic meaning of connections/influence.

Some other attested examples of infinitival relatives are the following:

- (78) a. Oh, Doni, there was never any *hatchet* to *bury*, you silly goose!
 (*La Di Da Di Bloody Da!* by Robin Anderson)
 b. I have a little *hell* to *raise*.
 (<http://ifunny.co/fun/zW6ZDE4R46>)
 c. So I *had* to actually socialize with the host's crew because we happen to sit in a small proximity and we got some *ice* to *break*.
 (<https://meganusa.wordpress.com>)

As described above, *bury the hatchet* can be decomposed into an NP meaning 'a quarrel' and a verb meaning 'put aside, lay to rest'. The NP *hatchet* has this meaning in both clauses in (78a), and hence it does not violate the CIC. In (78b), *raise hell* can be decomposed into an NP meaning 'trouble, mischief', while *raise* is fairly literal. 'Trouble, mischief' is again compatible with the verb *have*. In (78c), *break the ice* can be decomposed into an NP meaning 'stiff formality from lack of familiarity', and again *break* is fairly literal. The meaning of the NP is again compatible with the verb *have* (and we probably need a more detailed analysis of the pattern *have X to V*, which has a modal meaning rather than pure possession).

This fourth class of relative clause, with decomposable idioms, permits the NP of the idiom to serve as an argument of another verb and still be interpreted idiomatically. The NP serves as an argument in both the relative clause and the main clause. It can do this because the meaning that is assigned is not incompatible with the CIC.

Far from being problematic, then, these sorts of examples, like those in section 3.3, provide the strongest support for the CIC and against the movement view.

6.2.6 Actual Violations of the CIC

To summarize so far, there are four ways idioms are interpreted in relative clauses, listed below:

- (79) a. A silent *amount* interpretation;
 b. A propositional interpretation;
 c. The full idiom occurs in both the main and relative clause;
 d. The idiom is decomposable.

Only in (79c–d) does the NP actually serve as a thematic argument in both clauses. None of the four violate the CIC.

Now, when the idiom cannot be interpreted in one of these four ways, relativization does not work:

- (80) a. # John saw the *heed* that Mary *paid*.
 b. # I ate the *new leaf* that he *turned over*.
 c. # The *bucket* that James *kicked* tripped me up.
 d. # I polished the *plug* that he *pulled* on that project.
 e. # I fed the *dogs* that the police chief *called off*.

All of these violate the CIC, since the idiom chunk is assigned an interpretation in one position that is incompatible with its meaning in the idiom. On the movement view, it is not clear why these would be bad: the head noun should be able to get the idiomatic interpretation via the hypothesized movement chain.

Example (80a) is modeled after the following example:

- (81) * John saw the *heed* last year that Mary *paid*. (Henderson 2007, 215, (27a))

Henderson (2007) claims the problem here is extraposition of the relative clause, but as (80a) shows, the example is ungrammatical even without extraposition. Hulsey and Sauerland (2006) also claim that relative clauses with idiomatic heads cannot be extraposed, and according to them, this follows because they must have a head-raising analysis. But note that non-idiomatic relatives with the propositional interpretation discussed above also do not admit extraposition:

- (82) a. * They were shocked by the offers yesterday that I turned down.
b. * They couldn't believe the hypotheses last year that I rejected.

It appears that an NP+relative clause that is interpreted as a proposition does not permit extraposition. I will not attempt to explain this here, but note that this phenomenon is independent of idioms. Furthermore, as Henderson (2007) shows, when the idiomatic head occurs with the rest of the idiom in both clauses, extraposition is possible:

- (83) a. I *made* the same *headway* last year that he *made* this year.
b. We can *pull* the same *strings* this year that we *pulled* last year.
c. He is *barking up* the same *wrong tree* today that he was *barking up* yesterday.

This means that relative clauses with idiom chunks as heads can in principle extrapose.

Returning to the CIC, if an idiom chunk is incompatible with an 'amount' reading, an attempt at the silent amount type of relative clause fails:

- (84) a. # The soldiers *kicked* lots of *buckets*.
b. # The evildoer was pleased by the *buckets* that the soldiers *kicked*.
(85) a. # They *pulled* a lot of *plugs on* those projects.
b. # Tea Party Republicans were pleased by the *plugs* that they *pulled on* those projects.

Looking more closely at idiom chunks in relative clauses, then, we see that all acceptable examples satisfy the CIC, while examples that violate it are not well-formed. Again, what is important is the CIC, and movement is a non-issue.

6.2.7 Idioms in the Main Clause

One last issue with relative clauses is that an idiomatic head noun can also occur with the rest of its idiom just in the main clause, and not in the relative clause:

- (86) a. Bill *made* the amount of *headway* that Mary demanded. (Hulsey and Sauerland 2006, (47a))
b. Let's *cut* him the *slack* that he needs.
c. Pravda began to *toe a line* approved by Stalin and Kamenev.

Some of the examples of this that I have found involve the noun *amount*, either overtly or covertly. So the representation of (86a) is actually *John made the amount of headway that Mary demanded* <amount>, where what Mary demands is an amount. The idiom is confined to the matrix clause. In the other examples, though, an amount interpretation does not seem to be correct. These examples are also instances of decomposable idioms. *Cut X slack* includes an NP that means something like 'leeway' or 'allowances'. This is compatible with the verb *need* in the

embedded clause. *Toe the line* includes an NP that means something like ‘the approved actions’, with *toe* meaning ‘stick to’. ‘The approved actions’ is of course compatible with the verb *approve* in the embedded clause.

So, examples of idiom chunks occurring with the rest of the idiom just in the main clause and not in the embedded clause also fall into different categories, but they are the same categories that we saw with the converse pattern. All of them are explained by the CIC.

6.2.8 Summary

To summarize this section, idioms in relative clauses raise various issues, but closer inspection helps us to understand how they work. Relative clauses are not compatible with the view that the only dependency an idiom chunk can enter into is a movement dependency. Idiom chunks can actually occur in thematic argument positions that are not part of their idiom. All of the facts follow in the CIC analysis, and they are compatible with a relative operator analysis as well as the head raising analysis.

Again, it is not my purpose here to decide between the head raising analysis and the relative operator analysis, but it is worth pointing out that the expletive *there*, which does seem to be a valid movement diagnostic, is not grammatical as the head of a relative clause:

- (87) a. *I believe there [that occurred an accident] to have been a riot.
- b. *I think that there, which occurred an accident, was a riot.

This is surprising, because, in general, any two grammatical clauses that share an NP can be combined into one sentence where one of the clauses is a relative clause modifying the NP in the other clause:

- (88) a. I believe the woman to be a genius. The woman is wearing a hat. →
- b. I believe the woman(,) [who is wearing a hat](,) to be a genius.
- (89) a. I believe there to have been a riot. There occurred an accident. →
- b. *I believe there(,) [which/that occurred an accident](,) to have been a riot.

Expletives are an exception. It is not clear in the head raising analysis of relative clauses why they would be. One could appeal to semantics, and say that relatives restrict; then expletives are not the kind of thing that can be restricted. Non-restrictive relative clauses, though, cannot have expletives as heads, either (87b). These facts follow in the operator analysis, since, as we saw above, expletives are incapable of anteceding anything other than their own trace. They therefore cannot antecede a null operator. It is not at all clear if the head raising analysis can explain these facts.

7 A Theory of Idioms

In this section, I propose a theory of idioms that explains all of the facts we have seen in this paper. To begin, there is a vast literature on idioms, too vast to fully cite here (important landmarks include Fraser 1970, Nunberg, Sag, and Wasow 1994, Jackendoff 1997, Marantz 1997, Bruening 2010). As might be expected with such a vast literature, there are numerous accounts of idioms and where their special meaning comes from and is stored. These accounts can be divided into two broad approaches. One views idioms as being stored as wholes, with their special meaning listed as being associated with the whole. The Construction Grammar approach to idioms is emblematic of this view (e.g., Fillmore, Kay, and O’Connor 1988, Goldberg 1995), but versions of it can also be found in Jackendoff (1997, 2002) and Head-Driven Phrase Structure Grammar (e.g., Riehemann 2001). Such approaches simply stipulate idioms as exceptional, and they therefore predict much more idiosyncrasy than is actually found (e.g., Mateu and Espinal 2007). As much recent literature has observed (Marantz 1997, McGinnis 2002, Mateu and Espinal 2007), idioms are generally well-formed syntactic structures with certain aspects of their meaning following from their syntactic composition.

The other approach to idioms stores the special meaning of an idiom in the lexical entry of one or more of its component words or morphemes. For instance, Marantz (1997) views idioms as special meanings stored in

the lexical entries of roots. This special meaning is triggered when the root occurs in the local context of certain other elements. This view has to distribute idiomatic meanings across multiple lexical entries; for instance, *kick the bucket* has to list a special meaning for *kick* just in the context of *the bucket*, a special meaning for *the* just in the context of *kick* and *bucket*, and a special meaning for *bucket* just in the context of *kick* and *the* (Williams 2007). Other exemplars of this basic approach (e.g., Bresnan 1982, Sag and Wasow 2003) face similar problems, and others summarized in Riehemann 2001.

7.1 Proposal: Storage of Partial Syntactic Structures

I will propose a third, somewhat mixed view. Along with Marantz (1997), McGinnis (2002), Mateu and Espinal (2007), I view idioms as having a non-exceptional and largely compositional syntax. This is because the syntactic rules and constraints of the language are what they are and simply cannot be violated.⁶ At the same time, however, I view lexical entries in general as (partial) syntactic structures. As an example of what I have in mind, Bruening (2014) presents a theory of adjectival passives where an apparently non-phrasal adjectival passive like *proven* is actually an entire syntactic structure that includes a full VP and possibly even an embedded clause. In Bruening (2014), some NP (typically the object of the V) is a null operator that moves and abstracts over the entire VP to form an adjective that is a predicate of individuals (see Bruening 2014 for details). What is important here is that all words (and other stored items) are actually stored *phrases*, with specified pairings between syntactic structures and morphophonological items.

On this view, idioms are simply stored partial syntactic structures paired with some phonological content, exactly like words. There is nothing exceptional about phrasal idioms. *All* stored lexical entries are partial syntactic structures. In the limiting case, a stored lexical item might be a single syntactic head. This might be the case for a simple noun like *bush*, for example. Single words can also be complex phrases, however; for instance, the adjectival passive *bushed* ('exhausted') is a phrasal idiom with the full syntactic structure of an adjectival passive, as in Bruening (2014). It has an unexpected meaning that is stored with this syntactic structure (and note that there need not be any independently existing V *bush* from which the adjectival passive is built). It only differs from other phrasal idioms in corresponding only to a single morphophonological word rather than to multiple words. In theoretical terms, however, it is exactly the same as a phrasal idiom. Fixed expressions in general are exactly the same: stored pairings of syntactic structures with (morpho)phonological units. They differ from phrasal idioms only in special meaning: a phrasal idiom has a non-literal meaning that is stored with the stored pairing of syntax and phonology, but a literal fixed expression does not.

This view of idioms does not expect more idiosyncrasy than is observed, because all stored partial phrases must conform to the phrasal syntax. They have to, because they are in fact built and sanctioned by the phrasal syntax (the phrasal syntax is the only generative engine in this view; there is no generative lexicon). It also does not run into the problem of distributing the meaning of an idiom across multiple lexical entries. There is just one entry, for the idiom as a whole.⁷

Note that storing a phrasal idioms like *beat around the bush* as a partial syntactic structure is not the same as admitting constructions as a theoretical construct. There is no notion of a "construction" in this theory, since syntactic structures are not stored by themselves with particular meanings, in the absence of phonological forms. Instead, there are stored pairings of syntactic structures and phonological forms. Storing a meaning of 'obfuscate' with the partial syntactic structure associated with *beat around the bush* is doing nothing more than associating multiple meanings with a particular listed item. It is not different from listing the following meanings with the item *bush* associated with a partial NP syntactic structure: 'shrub'; 'wild or uncultivated country' (especially with *the*); 'thick growth of hair'; last name; etc. Phrasal idioms are exactly the same, with a meaning listed with the stored

⁶There are certain fixed expressions that do seem to violate syntactic rules, like *no can do* and *til death do us part*. I view these as being generated by a different grammar, for instance an archaic one whose word order is different or a grammar that imitates stereotypes of first- or second-language learners. I take language speakers to often have multiple grammars and to switch between them in different contexts (for instance, many English speakers switch between a grammar that does not allow negative concord and one that does).

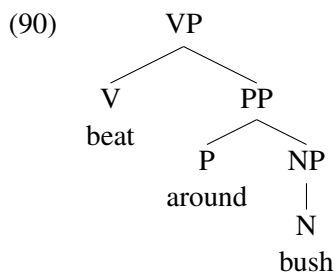
⁷Note that the lexicon in this conception is quite redundant: there is an entry for *beat around the bush*, an entry for *beat*, an entry for *bush*, another entry for *the bush* ('wild or uncultivated country'), etc. It is likely that speakers also store multiple tokens of the same word in phrasal contexts that they encounter, and extract commonalities from them.

phrase that is different from the compositional meaning.

Note also that storing syntactically complex objects is simply necessary for any complete theory of language. People store sayings (*you can't have your cake and eat it too*), collocations and other fixed expressions (*concerted effort, last but not least*), catch phrases, and even whole discourses in the form of jokes and stories. There is nothing exceptional about phrasal idioms.

So, all lexical entries on this view are partial syntactic structures paired with (morpho)phonological content. These partial syntactic structures are embedded within larger syntactic structures whenever they are put to use. These larger syntactic structures may trigger permutations of the partial substructure, such that inflection changes or phrases from within the partial substructure undergo displacement. The partial substructures can also themselves be altered to some extent by having modifiers and functional elements of various kinds added to them, in accordance with the syntax of the language.

To illustrate, the phrasal idiom *beat around the bush* is stored as a partial VP including the head V, a PP headed by *around*, and an NP headed by *bush* (I leave open whether verbs like *beat* and prepositions like *around* are themselves complex):



I assume that functional elements like determiners are typically not stored as part of the idiom, but are added following the syntactic rules of the language. They can take forms different from the canonical one, for instance (see also Riehemann 2001 and Bruening 2010):

- (91) Her friend wanted all the gory details and she wasn't going to *beat around **many** bushes* to get them.
(Michelle Garren Flye, *Where the Heart Lies*, accessed via Google Books)

So, when the above partial VP is included in a larger syntactic structure, functional elements within the partial VP will be specified, according to the syntactic context. In an inflection-rich language, this will include case and agreement markers in addition to determiners, and also tense and aspect, which we see in English (note also that *bush* becomes plural in (91)). As mentioned above, modifiers can also be added within the partial substructure (Ernst 1981, Nunberg, Sag, and Wasow 1994, Nicolas 1995, O'Grady 1998):

- (92) It is always a pleasure to hear Governor Brown answer questions directly and not *beat around the **political** bush*.
(<http://hereandnow.wbur.org/2014/03/18/california-drought-governor>)

There are of course constraints on what can be included in an idiom, and what can be added once the idiom is embedded within a larger syntactic structure. I assume these constraints are based on *selection*, as in Bruening (2010). See that work and Bruening (2015) for details. The elements that tend to vary are exactly the non-selected ones, like adjectives and possessors, while the idiomatic elements form a chain of selection (e.g., *beat* selects a PP headed by *around*, *around* selects an NP headed by *bush*). However, the exact nature of these constraints is not important here.

Before turning to anaphora, it should also be noted that some idioms permit some flexibility in the particular words that are involved. For example, *clutch at straws* is also attested as *grasp at straws*. I assume that the stored pairing of a partial syntactic structure with morphophonological items can include a finite list of such items instead of just a single item. So, the entry for *clutch/grasp at straws* might be exactly that, with a disjunction between *clutch* and *grasp*. Alternatively, the pairing might not include specific phonological forms, but might be a semantic and syntactic feature bundle instead, which can be associated with any item that matches (so, *clutch* and *grasp*

are both compatible with the syntactic and semantic specification of this particular idiom; we might also then find speakers using other related words, like *grab* and *clasp*, and indeed these are attested). Semantic classes are also relevant to “families of idioms” like *beat the crap out of*. Here the verb can be one of several verbs of physical violence, like *beat* and *kick*, and the noun can also vary within a limited range, as *crap*, *shit*, *tar*, etc. The result is that the particular words used in the idiom can even be somewhat productive.

7.2 Movement and Anaphora

With this theory in place, let us turn to movement and anaphora. I assume that, in order for a phrase that is being processed (in production or comprehension) to have a stored idiomatic meaning, it must match a stored partial syntactic structure and its pairing with morphophonology. In movement, this is simple. Consider raising:

(93) *The shit* seems to have *hit the fan*.

Adopting a movement approach to raising, *the shit* originated in the position of the external argument of *hit*. This means that this sentence includes the partial syntactic structure [the shit hit the fan], which matches what is stored with a special meaning. The special meaning is therefore made available by this matching.

Now consider a case of anaphora and ellipsis:

(94) When *the shit hit the fan*, and **it** did, we had a lot of shoulders to offload responsibility on.

(http://www.shattercolors.com/interviews/author_saknussemm.htm)

The adjunct clause here obviously matches the stored structure, so it has the idiomatic meaning. I assume that ellipsis is phonological deletion of full syntactic structure, so the phrase *and it did* includes an unpronounced VP *hit the fan*.⁸ This partially matches the stored idiom. The issue now is the pronoun *it*.

Research on the processing of pronouns has overwhelmingly concluded that they activate the lexical content of their antecedents. Studies find that both semantic aspects (Shillcock 1982) and phonological aspects (Schmitt, Meyer, and Levelt 1999) of the antecedent are accessed at the point of processing a pronoun, as are formal features (Cacciari, Carreiras, and Cionini 1997). The exact details of this activation do not matter here; all that matters is that pronouns do activate their antecedents in some manner. This means that in processing [it did ~~hit the fan~~], the pronoun *it* activates *the shit*. This is sufficient to lead to match for triggering the idiomatic meaning. I assume that the subject raised across the auxiliary from a VP-internal position (e.g., Koopman and Sportiche 1991), so there is a partial syntactic structure [it ~~hit the fan~~], where *it* is some form of pointer to its antecedent, *the shit*. I will represent this as a subscript on the pronoun, *it_{the shit}*. This leads to match: [it_{the shit} ~~hit the fan~~] matches the stored entry for the idiom, making the idiomatic interpretation available (the literal meaning will always be available, too, I assume).

I assume that the VP pro-form *do so* works similarly. Consider one of the examples from above:

(95) *The shit hits the fan* with astonishing regularity around here. How can it do so so frequently?

In the second sentence, we have the partial structure [it do so]. The pronoun *it* is some form of pointer to its antecedent *the shit*, as just described. I assume that the pro-form *do so* is similar, and points to its antecedent, *hits the fan*. The structure is therefore [it_{the shit} do so_{hit the fan}], with the pointers indicated. This structure matches the stored partial structure, and again the idiomatic meaning is available.

7.3 Control and Null Operators

Exactly the same mechanisms will be at play in cases of null anaphora, like control and hypothesized cases of null operators (if we accept null operator analyses). Consider control first:

(96) *The plug* is waiting to be *pulled* once all the 5600 stock is gone. . .

(<http://www.oesf.org/forum/lofiversion/index.php/t1554.html>)

⁸Psycholinguistic research on VP ellipsis finds phonological and semantic aspects of the antecedent VP being activated at the ellipsis site; see Tanenhaus, Carlson, and Seidenberg (1985), Shapiro and Hestvik (1995), Shapiro, Hestvik, Lesan, and Garcia (2003), and, for a good overview, Phillips and Parker (2014).

I assume that this has the following structure:

(97) The plug is waiting [PRO₁ to be [pulled *t*]]

The idiom here is *pull the plug (on)*, which I assume is stored as a partial VP with head *pull* selecting a complement NP headed by *plug*. The sentence above therefore has to find a match for this stored partial VP.

I assume that PRO is just like an overt pronoun in activating its antecedent. That is, it serves as a pointer to its antecedent, as described above. PRO in this example is the subject of a passive, so I assume that it moved from object position. Traces of movement also activate their antecedents, so, through a chain of antecedence, we end up with the VP [pulled PRO_{the plug}]. This matches the stored entry, and makes the idiomatic meaning stored with that entry available.

Note that the overt NP, *the plug*, is a thematic argument of *wait*. This does not interfere with accessing the idiomatic meaning in any way, since, as argued here, the only constraint is the CIC. *The plug* cannot be assigned some meaning that is incompatible with the meaning it requires for the idiom. As described above, *wait* imposes no restrictions on its subject, so the CIC is satisfied.

Consider now a parasitic gap case, assuming the null operator analysis of parasitic gaps:

(98) We were amazed at the *strings*₁ she managed to *pull* *t*₁ [OP₂ without seeming to *pull* *t*₂].

Here, in the matrix clause the trace of *strings* is a pointer to its antecedent, so we end up with the VP [pull *t*_{strings}], which matches the stored entry for the idiom. In the adjunct clause, the trace in object position is the trace of the null operator. Some mechanism must combine the null operator chain with the main clause chain (Chomsky 1986, Nissenbaum 1998). This makes the ultimate antecedent for the embedded trace also *strings*, so we again have a VP [pull *t*_{strings}]. This matches the stored entry, and again the idiomatic meaning is available.

7.4 Extrasentential Anaphora

As described in section 3.3 above, there are also attested cases like the following:

(99) They had the *strings* and *pulled* them. (Riehemann 2001, 75, (73e))

In such cases, part of the idiom appears as a thematic argument of a completely different predicate, while only the anaphor occurs with the rest of the idiom.

These appeared to be problematic, but as described in section 3.3 actually turned out to provide important support for the point of view outlined here. They also now follow in the theory being proposed. In the second conjunct, we have a VP [pulled them_{the strings}], with the pronoun serving as a pointer to its antecedent, as described above. This matches the stored entry for [pull strings], and again the idiomatic meaning is available. It does not matter that the antecedent for the anaphor is the thematic argument of some other predicate, so long as the CIC is respected. In this case, it is; the verb *have* does not impose any particular requirements on its object, and so the first conjunct is compatible with the meaning of the idiom, where *strings* is metaphorical for ‘connections’ or ‘influence’.

Now, importantly, it was crucial for accessing the idiomatic meaning that the entire partial structure of the idiom be matched. This then predicts that even though *pull strings* can be decomposed, neither part can occur on its own. This appears to be correct:

(100) a. # They had the strings.
b. # Pull!

Neither *strings* nor *pull* match the stored partial structure [pull strings], so the idiomatic meaning cannot be accessed.

The same is true with the examples of infinitival relative clauses that we saw in section 6.2.5:

(101) a. Why would a reputed master politician fail to find one *string* to *pull*? (Riehemann 2001, 75, (73a))

- b. # Why would a reputed master politician fail to find one *string*?
- (102) a. But Lee has even more *strings* to *pull* . . . (Riehemann 2001, 76, (73g))
- b. # But Lee has even more *strings*.

This distinguishes the theory proposed here from one that simply treats decomposable idioms as combinations of metaphorical constituents. Such a view would predict that the parts could occur on their own, but this does not generally seem to be true.⁹

7.5 Summary

This section has proposed a new theory of idioms where they are stored pairings of partial syntactic structures with morphophonological content. This is exactly what words are, too, as well as fixed expressions, collocations, etc. Idiomatic meaning is accessed when part of a syntactic structure matches the stored partial structure. Anaphoric processes can lead to match because anaphors point to the content of their antecedents. This theory predicts exactly the range of facts described in this paper.

8 Conclusion

This paper has shown that idioms are not a valid movement diagnostic. Idiom chunks can participate in control and pronominal anaphora, and they must be able to antecede null operators in null operator analyses of various phenomena. Idiom chunks can also appear in thematic argument positions of predicates other than the rest of their idiom, even across sentence boundaries. The only constraint they must obey is the CIC, which itself is simply a variant of the Anti-Pun Ordinance of Zaenen and Karttunen (1984). An element cannot be assigned two different interpretations at the same time.¹⁰

This paper also proposed a new analysis of idioms (and lexical items generally), where all lexical entries are partial syntactic structures paired with morphophonological items. Idiomatic meanings are triggered when a part of a syntactic structure being processed matches the stored partial structure-form pair. This theory explains all of the data presented in this paper, and successfully accounts for the large but not unconstrained flexibility that idioms have been shown to exhibit. In particular, it accounts for cases where idiom chunks occur as arguments of predicates other than the rest of their idiom.

References

- Akmajian, Adrian, and Frank Heny (1975), *An Introduction to the Principles of Transformational Syntax*. Cambridge, MA: MIT Press.
- Berman, Arlene (1973), "A Constraint on Tough-Movement." In Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser, eds., *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, Chicago: University of Chicago, pp. 34–43.
- Bianchi, Valentina (1999), *Consequences of Antisymmetry: Headed Relative Clauses*. Berlin: Mouton de Gruyter.

⁹There are some cases where a part of an idiom becomes used by itself, for instance *bucket* in the new idiom *bucket list* (a list of things you want to do before you kick the bucket). I assume that this is a metalinguistic process of creating new words and phrases by reference to others.

¹⁰Of course, one can violate the Anti-Pun Ordinance in order to make a pun. This happens with idioms, for instance in the following attested examples:

- (i) But like 19th-Century Populist Mary Elizabeth Lease, who exhorted Kansas farmers to "raise less corn and more hell," they do get riled up with great regularity. (Riehemann 2001, 80, (80e))
- (ii) Would Mags and Raylan bury the hatchet? Or bury it in each other's backs?
(<http://www.televisionaryblog.com/2011/05/put-end-to-my-troubles-getting-to-know.html>)

Both of these examples violate the CIC, but they do so deliberately, as they are intentional puns.

- Bouton, Lawrence (1970), "Antecedent Contained Proforms." In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, Chicago: University of Chicago.
- Bresnan, Joan (1971), "Note on the Notion 'Identity of Sense Anaphora'." *Linguistic Inquiry* 2: 589–597.
- Bresnan, Joan (1982), "The Passive in Lexical Theory." In Joan Bresnan, ed., *The Mental Representation of Grammatical Relations*, Cambridge, MA: MIT Press, pp. 3–86.
- Brody, Michael (1993), " θ -Theory and Arguments." *Linguistic Inquiry* 24: 1–23.
- Bruening, Benjamin (2010), "Ditransitive Asymmetries and a Theory of Idiom Formation." *Linguistic Inquiry* 41: 519–562.
- Bruening, Benjamin (2014), "Word Formation is Syntactic: Adjectival Passives in English." *Natural Language and Linguistic Theory* 32: 363–422.
- Bruening, Benjamin (2015), "Idioms and Locality." Ms., University of Delaware, available at <http://udel.edu/~bruening/Downloads/IdiomsLocality1.pdf>.
- Cacciari, Cristina, Manuel Carreiras, and Cristina Barbolini Cionini (1997), "When Words Have Two Genders: Anaphor Resolution for Italian Functionally Ambiguous Words." *Journal of Memory and Language* 37: 517–532.
- Chomsky, Noam (1977), "On WH-Movement." In Peter Culicover, Thomas Wasow, and Adrian Akmajian, eds., *Formal Syntax*, New York: Academic Press, pp. 71–132.
- Chomsky, Noam (1981), *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1986), *Barriers*. Cambridge, MA: MIT Press.
- Davies, William D., and Stanley Dubinsky (2004), *The Grammar of Raising and Control*. Oxford: Blackwell.
- Ernst, Thomas (1981), "Grist for the Linguistic Mill: Idioms and "Extra" Adjectives." *Journal of Linguistic Research* 1: 51–68.
- Evans, Gareth (1980), "Pronouns." *Linguistic Inquiry* 11: 337–362.
- Fiengo, Robert, and Robert May (1994), *Indices and Identity*. Cambridge, MA: MIT Press.
- Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor (1988), "Regularity and Idiomaticity in Grammatical Constructions." *Language* 64: 501–538.
- Fraser, Bruce (1970), "Idioms within a Transformational Grammar." *Foundations of Language* 6: 22–42.
- Goldberg, Adele E. (1995), *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Haegeman, Liliane (1994), *Introduction to Government and Binding Theory*. Oxford: Blackwell, second ed.
- Hallman, Peter (2013), "Predication and Movement in Passive." *Lingua* 125: 76–94.
- Hankamer, Jorge, and Ivan Sag (1976), "Deep and Surface Anaphora." *Linguistic Inquiry* 7: 391–426.
- Hartman, Jeremy (2009), "Intervention in Tough Constructions." In Anisa Walkow, Martin Schardl, and Muhammad Abdurrahman, eds., *Proceedings of the 39th Meeting of the North East Linguistic Society*, Amherst, MA: GLSA, University of Massachusetts, p. to appear.
- Hartman, Jeremy (2012), "(Non-)Intervention in A-Movement: Some Cross-Constructional and Cross-Linguistic Considerations." *Linguistic Variation* 11: 121–148.
- Henderson, Brent (2007), "Matching and Raising Unified." *Lingua* 117: 202–220.
- Hicks, Glyn (2003), *So Easy to Look At, So Hard to Define: Tough Movement in the Minimalist Framework*. Master's thesis, University of York.
- Hicks, Glyn (2009), "Tough-Constructions and Their Derivation." *Linguistic Inquiry* 40: 535–566.
- Hornstein, Norbert (1999), "Movement and Control." *Linguistic Inquiry* 30: 69–96.
- Hornstein, Norbert (2001), *Move! A Minimalist Theory of Construal*. Cambridge, MA: Blackwell.
- Hulsey, Sarah, and Uli Sauerland (2006), "'Sorting Out Relative Clauses.'" *Natural Language Semantics* 14: 111–137.
- Jackendoff, Ray S. (1997), *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Jackendoff, Ray S. (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford and New York: Oxford University Press.
- Karttunen, Lauri (1969), "Pronouns and Variables." In Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, eds., *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*, Chicago: University of Chicago, Chicago Linguistic Society, pp. 108–115.

- Kayne, Richard (1994), *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Koopman, Hilda, and Dominique Sportiche (1991), "The Position of Subjects." *Lingua* 85: 211–258.
- 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, and Alexander Williams, eds., *Proceedings of the 21st Annual Penn Linguistics Colloquium*, no. 4:2 in Penn Working Papers in Linguistics, Philadelphia: University of Pennsylvania.
- Mateu, Jaume, and M. Teresa Espinal (2007), "Argument Structure and Compositionality in Idiomatic Constructions." *The Linguistic Review* 24: 33–59.
- McGinnis, Martha (2002), "On the Systematic Aspect of Idioms." *Linguistic Inquiry* 33: 665–672.
- Merchant, Jason (2001), *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.
- Munn, Alan (1993), *Topics in the Syntax and Semantics of Coordinate Structures*. Ph.D. thesis, University of Maryland College Park.
- Nicolas, Tim (1995), "Semantics of Idiom Modification." In Martin Everaert, Erik-Jan van der Linden, André Schenk, and Rob Schreuder, eds., *Idioms: Structural and Psychological Perspectives*, Hillsdale, NJ: Lawrence Erlbaum, pp. 233–252.
- Nissenbaum, Jonathan (1998), "Movement and Derived Predicates: Evidence from Parasitic Gaps." In Uli Sauerland and Orin Percus, eds., *The Interpretive Tract*, Cambridge, MA: MITWPL, vol. 25 of *MIT Working Papers in Linguistics*, pp. 247–295.
- Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow (1994), "Idioms." *Language* 70: 491–538.
- Nunes, Jairo (2004), *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
- O'Grady, William (1998), "The Syntax of Idioms." *Natural Language and Linguistic Theory* 16: 279–312.
- Pesetsky, David (2012), "Phrasal Movement and Its Discontents: Diseases and Diagnostics." In Lisa Cheng and Norbert Corver, eds., *Diagnostics in Syntax*, Oxford: Oxford University Press, p. to appear.
- Phillips, Colin, and Dan Parker (2014), "The Psycholinguistics of Ellipsis." *Lingua* 151: 78–95.
- Postal, Paul M. (1971), *Cross-Over Phenomena*. New York: Holt, Rinehart, and Winston.
- Postal, Paul M. (1974), *On Raising: One Rule of English Grammar and Its Theoretical Implications*. Cambridge, MA: MIT Press.
- Postal, Paul M. (2004), *Skeptical Linguistic Essays*. Oxford: Oxford University Press.
- Riehemann, Susanne Z. (2001), *A Constructional Approach to Idioms and Word Formation*. Ph.D. thesis, Stanford University.
- Ruwet, Nicolas (1991), "On the Use and Abuse of Idioms in Syntactic Argumentation." In Nicolas Ruwet and John A. Goldsmith, eds., *Syntax and Human Experience*, Chicago: University of Chicago Press, pp. 171–251.
- Sag, Ivan (1976), *Deletion and Logical Form*. Ph.D. thesis, Massachusetts Institute of Technology. Distributed by MIT Working Papers in Linguistics, Cambridge, Mass.
- Sag, Ivan, and Thomas Wasow (2003), *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications.
- Sauerland, Uli (1998), *The Meaning of Chains*. Ph.D. thesis, Massachusetts Institute of Technology. Distributed by MIT Working Papers in Linguistics, Cambridge, Mass.
- Schachter, Paul (1973), "Focus and Relativization." *Language* 49: 19–46.
- Schmitt, Bernadette M., Antje S. Meyer, and Willem J. M. Levelt (1999), "Lexical Access in the Production of Pronouns." *Cognition* 69: 313–335.
- Shapiro, Lewis P., and Arild Hestvik (1995), "On-Line Comprehension of VP-Ellipsis: Syntactic Reconstruction and Semantic Influence." *Journal of Psycholinguistic Research* 24: 517–532.
- Shapiro, Lewis P., Arild Hestvik, L. Lesan, and A. R. Garcia (2003), "Charting the Time-Course of VP-Ellipsis Sentence Comprehension: Evidence for an Initial and Independent Structural Analysis." *Journal of Memory and Language* 49: 1–19.
- Shillcock, R. (1982), "The On-Line Resolution of Pronominal Anaphora." *Language and Speech* 25: 385–401.
- Sudo, Yasutada (2015), "Hidden Nominal Structures in Japanese Clausal Comparatives." *Journal of East Asian Linguistics* 24: 1–51.

- Tanenhaus, Michael K., Gregory N. Carlson, and M.S. Seidenberg (1985), "Do Listeners Compute Linguistic Representations?" In David R. Dowty, Lauri Karttunen, and Arnold M. Zwicky, eds., *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, Cambridge: Cambridge University Press, pp. 359–408.
- Vergnaud, Jean-Roger (1974), *French Relative Clauses*. Ph.D. thesis, Massachusetts Institute of Technology.
- Williams, Edwin (2007), "Dumping Lexicalism." In Gillian Ramchand and Charles Reiss, eds., *The Oxford Handbook of Linguistic Interfaces*, Oxford: Oxford University Press, pp. 353–381.
- Zaenen, Annie, and Lauri Karttunen (1984), "Morphological Non-Distinctiveness and Coordination." *Proceedings of the Eastern States Conference on Linguistics* 1984: 309–320.
- Zhang, Niina Ning (2010), *Coordination in Syntax*. Cambridge: Cambridge University Press.