Against the parallelism between the NP and the clause: Evidence from idioms

Carlo Cecchetto, Caterina Donati

This a prefinal version of the following paper:

AGAINST THE PARALLELISM BETWEEN THE NP AND THE CLAUSE: EVIDENCE FROM IDIOMS

Carlo Cecchetto
SFL (CNRS & Paris VIII) and
University of Milan-Bicocca

Caterina Donati
LLF, Université Paris Diderot

carlo.cecchetto123@gmail.com

cdonati@linguist.univ-paris-diderot.fr

Abstract

We argue that the counterpart of Marantz's generalization does not hold in the nominal domain, because there are idioms in which the determiner and the noun receive an idiomatic meaning while the PP that modifies the noun is not part of the idiom (we call these "PP-less idioms"). We show that PP-less idioms are fully expected if the hypothesis of parallelism between nominal structure and clausal structure is dropped and it is assumed that the first step of the derivation in the nominal domain involves merge of D and N.
As for the mirror image of PP-less idioms, "PP-containing idioms", namely DPs where N and the PP that follows the noun receive an idiomatic reading while D does not, we suggest that they are not generated by syntax but are rather the output of the morphological component.

1. Introduction

The topic of this paper is idioms and what they can tell us about the internal structure of the nominal domain. More specifically, we will identify a special class of idioms that have been neglected in the literature so far and argue that a proper analysis of these idioms requires departing from approaches that posit a close parallelism between nominal structure and clausal structure.

We start by discussing some pieces of evidence against the hypothesis that nouns take complements like verbs do (Section 2). In Section 3, we switch to the syntactic conditions under which an idiomatic reading may arise and defend the traditional claim that a category must form a constituent in order to be able to receive an idiomatic reading. This theory is built to capture the observation (often called "Marantz's generalization") that no idiom exists that includes the subject and the verb but not the object, while many idioms exist that include the verb and the object but not the subject.

In Section 4 we offer our main empirical contribution. We show that the counterpart of Marantz's generalization does not hold in the nominal domain, because there are idioms in which the determiner and the noun receive an idiomatic meaning while the PP that modifies the noun is not part of the idiom (we call these "PP-less idioms"). We show that PP-less idioms are fully expected if the hypothesis of parallelism between nominal structure and clausal structure is dropped and it is assumed that the first step of the derivation in the nominal domain involves merge of D and N.

Section 5 extends the discussion of idioms in the nominal domain, notably by considering idioms that are the mirror image of PP-less idioms, namely DPs where N and the PP that follows the noun receive an idiomatic reading while D does not. We call these "PP-containing idioms". We discuss several pieces of evidence that suggest that while PP-less idioms are formed in syntax, PP-containing idioms are the output of the morphological component. In Section 6, we consider whether PP-less idioms can be analysed as measure phrases or involve inverted predication. Section 7 concludes the paper.

2. Do nouns take complements?

In the generative tradition an influential line of research stresses the structural parallelism between the clausal and nominal domain. For example, it has been proposed that the nominal domain is divided into lexical and functional subdomains in a fashion similar to the clausal domain. A classical work in this tradition is Abney's (1987) DP hypothesis, which postulates a layer on the top of the nominal domain which is parallel to the CP layer postulated for the clausal domain. In later work it has been proposed that the functional projections corresponding to agreement and tense parallel dedicated functional projections internal to DP (cf. Bernstein 2008 for an overview).

In this paper we will be concerned with another facet of the purported parallelism between nominal and clausal domains, namely the hypothesis, stemming from Chomsky (1970), that at least some nouns, deverbal (*destruction*) and relational (*mother*) ones being the paradigmatic cases, take complements as verbs do. Although this property of nouns is linked to their semantics (for example, the event of destruction requires that something be destroyed and the concept of mother involves as a relatum the concept of son/daughter), the proposal is about syntax, since a slot in the functional projection of the noun is postulated.

The hypothesis that nouns take complements has not gone unchallenged (cf. Dowty 2003, Higginbotham 1985, Hale and Keyser 2002, Kayne 2009) and in this paper we want to add further reasons to reject it, following Adger (2013) and Cecchetto & Donati (2015) who recently proposed articulated arguments to this effect.

A first reason of suspicion is that complements of genuine transitive verbs are obligatorily expressed (cf. *"They often destroy…"). However, the purported complement of the noun is always dispensable ("Italian mothers are always anxious").[1] Therefore, in order to maintain the hypothesis that nouns take complements, it is necessary to make the *ad hoc* stipulation that the Theta-Criterion behaves differently depending on whether it applies to nouns or verbs.

Another reason to doubt the hypothesis that nouns take complements has to do with constituency tests. The contrast between (1) and (2) supports the view that verb and object form a minimal constituent that excludes the subject. In particular, the strong deviance of (2) is standardly explained by the fact that the unit formed by subject+verb excluding the internal argument can never be a constituent.

(1) *John bought a house and Mary did that too.*
(2) **John bought a house but did that no car*
     (Intended meaning: John bought a house but not a car)

However, if we apply the same type of constituency test to the unit formed by the noun and its alleged complement, the results are quite different. A proform like 'that' can replace the

---

[1] But see Grimshaw (1990) for a possible exception involving complex-event-nominals.

unit formed by determiner + noun, crucially excluding the alleged complement of the noun (cf. 3).

(3) *I have already seen the picture of John, but I haven't yet seen that of Mary.*

Converging evidence comes from another diagnostic of constituency: the movement test. Although English does not allow left-branch extraction, the unit formed by *wh*-determiner and noun can be fronted while the alleged complement of the noun is stranded,

(4) *Which picture did you see of me?*

This suggests that determiner and noun form a unit that excludes the rest of the NP/DP. Others reasons to question the parallelism between verb complementation and noun complementation have been proposed in the literature but it is time to move to the argument based on idioms.

3. Idioms in the verbal domain

In this section we first offer a working definition of idioms; we then present a well-known property of idioms in the verbal domain, which, following the previous literature, we call Marantz's generalization, and we defend it from some possible criticism. This section sets the stage for Section 4, where we show that idioms in the nominal domain do not obey Marantz's generalization.

3.1 A working definition of idiomatic expressions

Following a consolidated tradition, we take idioms to be expressions that are not derived by *normal* compositional processes from their parts and have a conventional meaning. The word 'normal' is crucial here, because, as Wasow, Sag & Nunberg (1994) have stressed, there are idioms that are somehow compositional, but are still conventional in their meaning. This is at the basis of the difference between idioms like *spill the beans* or *pull strings* on the one hand and idioms like *kick the bucket* or *shoot the breeze* on the other hand. The meaning of the former two idioms can be indeed traced back to the meaning of their components, while this is much harder for the latter two. Still, even if the meaning of the phrase *pull strings* can be derived from the meaning of *pull* and *strings* once one knows the idiomatic meaning, the composition process underlying idiom formation is not the standard composition process which leads to the literal reading.

Since their meaning is never the result of the standard compositional procedure, idioms must be listed as such in the mental lexicon, although *how* they are stored is open to discussion. For the time being, it is important to be clear on a point: the fact that an idiom must be stored in the lexicon does not imply that it must be assimilated to a non-syntactic object, namely a word or whatever output the morphology delivers to syntax (assuming that a separation between morphology and syntax can be drawn). For example, Di Sciullo & Williams (1987) and Williams (2007) have argued that an object listed in the lexicon is any kind of meaning-form pair that, since it cannot be derived by the application of compositional rules, needs to be learnt by heart by speakers. As this point will be important later in the paper, we will go back to it.

3.2 The cyclic theory of idiom formation and some challenges

What expressions can get an idiomatic reading? We assume without discussion that words can get an idiomatic meaning next to a literal one. The trickiest question is under what conditions complex expressions (phrases) can become idioms. A classical approach is a cyclic theory of idiom formation that goes back to Marantz (1984), which is our focus in this section.

3.2.1 Marantz's generalization

According to Marantz (1984), a category must form a constituent at some point of the derivation in order to receive an idiomatic reading. In the terminology of the Government and Binding theory, this was expressed by saying that all the idiom parts must be in a local configuration at D-structure. In Minimalist terms, this can be restated by saying that all the parts of an idiom must be contained in a single interpretive domain (a phase, cf. Harley and Stone 2013).
An observation that led Marantz to adopt this theory is that the exact meaning of a verb depends on the denotation of its internal argument. For example, although a core semantic nucleus is shared in all the sentences in (5), the meaning of the verb *kill* varies in a way that shows that its polysemy is dependent on the object that follows it.

(5)   a. *He killed a bug* = He caused the bug to die
         b. *He killed a conversation* = He caused the conversation to end
         c. *He killed an evening* = He whiled away the time span of the evening
         d. *He killed a bottle* = He emptied a bottle
         e. *He killed an audience* = He entertained the audience to an extreme degree

This approach allows for a kind of 'continuum' of idiomaticity, according to which idioms are subject to the same kind of interpretive process that a verb like *kill* is subject to. What changes is that the meaning of the verb in an idiom is dependent on a particular DP while the meaning of 'kill' in (5) is dependent on a class of DPs.
While this co-dependence in meaning in which idioms are the extreme case holds true for verb and object, Marantz (1984) proposed that no idiom exists that includes the subject and the verb but not the object. Under the cyclic theory of idiom formation, this is explained because subject and verb do not form a constituent that excludes the direct object at any point of the derivation. We will call "Marantz's Generalization" the claim that idioms in the clausal domain involve object + verb to the exclusion of the subject but never involve subject + verb to the exclusion of the object.
Marantz's Generalization can be seen as part of a more general condition, which is called Verb-Object Constraint by Baker (2009). This constraint states that verb and object form a minimal constituent that excludes the subject, and, as discussed by Baker, it is one of the best candidates for a language universal, since it applies also to polysynthetic languages that *prima facie* appear to be non-configurational.
Marantz's generalization has not been unchallenged. For simplicity, we discuss idioms that are potential counterexamples in different sub-sections.

3.2.2 Subject + V idioms?

The idioms in (6) are a *prima facie* counterexample to Marantz's generalization, since the subject and the verb are fixed and receive an idiomatic reading while the direct object is variable.

(6)   a. [*Lady Luck smile on X*] (Horvath and Siloni 2002)
     b. [*The vultures are circling X*] (Manaster-Ramer, *Linguist List* January, 28 1993)
     c. [*A little bird tell X Y*] (Wasow, Sag & Nunberg 1996)

However, it is far from obvious that the sentences in (6) are real counterexamples. Harley and Stone (2013) suggest that neither (6a) nor (6b) qualify as idiomatic subject + verb combinations, because 'Lady Luck' and 'the vultures' are DP idioms (like *The Big Apple*). Both these DPs can occur with other verbs in other positions, and still receive the same interpretation: *I hate those payday loan vultures; He has never once been visited by Lady Luck, Lady Luck laughed in his face.* In a similar vein Bruening (2010) claims that 'a little bird' in (6c) is a DP idiom, too. This is supported by examples such as *A little bird whispered/emailed it to me*, *A little bird is broadcasting that*, and *I heard it from a little bird*. If the idiomatic category is just the subject DP, the sentences in (6) are not counterexamples to Marantz's generalization.

3.2.3 Idioms with dislocated objects

In Italian it is possible to find idioms with a fixed subject + verb and a variable object in cases like the following:

(7)   *Lui/Gianni/Piero, se   lo   mangieranno i   vermi*
     He/Gianni/Piero SELF   him   will-eat   the   worms
     'He/Gianni/Piero will die'

In (7) the object is left dislocated and the subject is postverbal. Crucially, the idiomatic reading is not available if the object DP appears in the canonical postverbal position, no matter if the subject is preverbal (8) or postverbal (9).

(8)   *I   vermi si   mangieranno   lui/Gianni/Piero*
     the worms SELF will-eat   him/Gianni/Piero
     #"He/Gianni/Piero will die"
     'Worms will devour him/Gianni/Piero'

(9)   *Si   mangieranno   lui/Gianni/Piero, i vermi*
     SELF will-eat   him/Gianni/Piero the worms
     #"He/Gianni/Piero will die"
     'Worms will devour him/Gianni/Piero'

We conclude that the idiomatic reading is possible only if the object is not VP internal. Although we cannot provide a full analysis, we observe that this is expected if the only way

to assign an idiomatic reading to subject and verb is to idiomatize the entire *v*P constituent. A possible analysis of these cases is that the *v*P is given the idiomatic reading after the object has been left dislocated. In this way, the object is immune to the idiomatization process without this violating Marantz's generalization.[2]

3.2.4 The game "Simon says"

An indirect yet intriguing potential counterexample to Marantz's generalization is a game which in English is called "Simon says…." and in French is called "Jacques a dit…" .

In this game a player takes the role of Simon or Jacques and issues instructions such as "close your eyes" or "stick out your tongue" to the other players. These orders should only be obeyed if preceded by the words 'Simon says….' or 'Jacques a dit…'.

The command in this game looks like a counterexample to Marantz's generalization, since the subject and verb are fixed and formulaic, while X (the action to be performed) is variable. Notice however that X cannot be an embedded clause or a DP. If someone says *Jacques a dit que vous devez vous lever / Simon says that you have to stand up*, people don't have to stand up according to the rules of the game. They should obey only if the instruction is given as a direct order (*Jacques a dit: Levez vous!/Simon says: Stand up!*). This specific restriction would not follow if the command were a subject + verb idiom with an open slot for the object. A possible alternative analysis is that the fixed formula includes a null object, anaphoric to the following imperative sentence ("Simon said this: stand up"). Given this analysis the entire sentence, including the null object, is the idiom, in compliance with Marantz's generalization.

In this section we have introduced Marantz's generalization and we have argued that it is valid in the clausal domain, despite some challenges that have been proposed.[3]

4. Unexpected idioms in the nominal domain: PP-less idioms and their analysis

---

[2] A similar case is the Spanish idiom in (i), pointed out to us by Ane Berro. As shown in (ii), only the weird cannibalic reading is acceptable if the sentence has the SVO order.
(i)   *A Juan/Pedro/Jorge   le   come   Ramón*
        to Juan/Pedro/Jorge him eats   Ramon
'Juan/Pedro/Jorge is very bored'
(ii)   *Ramon (le)   come   a Juan/Pedro/Jorge*
        Ramon (him)   eats   to Juan/Pedro/Jorge
#" Juan/Pedro/Jorge is very bored"
'Ramon is eating Juan/Pedro/Jorge'

[3] Chtareva (2004), as reported by Harley and Stone (2013), discusses some interesting idioms in Russian that are another challenge to Marantz's generalization. One of them is reported in (i). (*Kondrashka* is a personal male name which in its idiomatic use refers to paralysis).
 (i)   Ivana   chut'   KONDRASHKA   ne   (s)xvat-il.
        Ivan-ACC   almost Kondrashka-NOM   not   grabbed
'Ivan was almost grabbed by paralysis' = 'Ivan was frightened to death'.
Chtareva argues that the sentence in (i) exhibits the syntactic behavior associated to Russian psych-predicates, rather than the behavior typically associated with agentive verbs. If the subject of object experiencer predicates is base-generated VP-internally (Belletti and Rizzi 1988, Pesetsky 1995), a sentence like (i) is indeed compatible with Marantz's generalization, since the surface subject is base-generated as an internal argument.

In this section we argue that the counterpart of Marantz's generalization is not valid in the nominal domain. In order to do that, we show that idioms exist in which the determiner and the noun receive an idiomatic meaning while the PP selected by the noun is not part of the idiom. For concreteness we will call these structures, which are our main concern in this paper, "PP-less idioms".

A first type of PP-less idiom is illustrated by the Italian expression *alla faccia di* (lit. to-the face of), which indicates that an action is performed or an event takes place which has the consequence that someone, who often is in a positon of authority, will be astonished/thrilled:

(10) *Alla faccia di tuo marito/del professore/del tuo capo* etc.
     To-the face of your husband/ of-the professor/your boss etc.
     'So much for your husband/ our professor/your boss'

Another example of the same type of PP-less idiom involves the expressions *uno straccio di* (lit. a rug of), a minimizer.[4]

(11) *Uno straccio di prova /marito/ lavoro*
     a rug of evidence/husband/job
     'a very poor job/evidence/husband'

The use of this expression is exemplified in the following sentences.

(12) *Mi basterebbe uno straccio di prova*
     To-me would-suffice a rug of evidence
     'A small evidence would be enough'

(13) *Si è ritrovata con uno straccio di marito*
     (she) herself is found with a rug of husband
     'She ended up with a lousy husband'

(14) *Non sono riuscito a trovare uno straccio di lavoro*
     (I) NEG have been-able to find a rug of work
     'I could not find any job at all'

Still another example of the first type of PP-less idioms is the expression *un salto* (lit. a jump), meaning a brief visit.

(15) *Un salto (in pizzeria / al mare / in vespa,...)*
     A jump (to pizzeria / to-the sea / in vespa)
     'A brief visit to a pizzeria, to the seaside, by driving my Vespa'

In order to be sure that we are dealing with genuine cases of PP-less idioms in the intended sense we need to make sure that
(i) the noun is part of the idiom, having lost its original meaning
(ii) the D is fixed.
Point *(ii)* is very important because, if the D were not fixed, we might be facing a case of noun polysemy, therefore these examples would tell us nothing about the extension of Marantz's generalization to the nominal domain. In fact, the D *is* fixed. If we replace *uno*

---

[4] See section 7 for a qualification about this idiom.

("a") with *lo* ("the") in (10), the idiomatic meaning is totally lost and expression "to a face of…" can only be used literally to talk about a fictional entity that has two faces.

Similarly, if we replace *uno* ("a") with *lo* ("the") in (11) this expression becomes gibberish, much as the English counterpart "the rug of job" or "the rug of evidence". Finally, if we replace *uno* ("a") with *lo* ("the") in (15), only the literal meaning surfaces *(il salto in pizzeria* means the jump performed inside a pizzeria). Interestingly, changing the determiner in this latter case can even give rise to a *different* PP-less idiom, as is the case with *quattro salti* (lit. "four jumps"), which means "to dance".

A second type of PP-less idioms involves a structure bigger than the DP, namely cases in which a verb, a determiner and a noun are part of an idiom while the PP following the noun is not.[5]

(16) *Essere a un passo dalla vittoria/rovina/fama*
      To-be at one step from victory/disaster/fame
      'To be very close to victory/disaster/fame'

The following idioms in French illustrate the same pattern:[6]

(17) *Etre à deux doigts de ...*
      To-be at two fingers of…
      'To be very close to…'

 (18) *Ne pas toucher un cheveux de la tête de ...*   (French)

---

[5] A reviewer challenges the idiomatic nature of (16) based on the fact that a similar idiom exists in English and other languages ("A studio located at two steps from the beach/village center"). We do not concur with this objection. The fact that many languages have the same or a similar idiom might be due to the fact that the path to idiomatization is more direct with some idioms, probably because the idiom is transparent. Crucial evidence that we are dealing with an idiom in (16) is that the determiner is frozen. For example, In Italian the expression *essere a due passi da…* ("to-be at two steps from") indicates only physical distance and cannot be used with the same range of meaning attested in in (13).

[6] Another case of PP-less idiom in French might 'Il n'y a pas trente-six façons de', illustrated in (i),

(i) *Il n'y a pas trente-six façons de le dire/de le faire/ de…*
      there aren't thirty-six ways of it saying /of it doing /of…
      'There is only one way to do/say/…'

This idiom, which roughly means "there is only one way to (do something)" excludes its PP complement. This is confirmed by the fact that in principle any complement can follow 'façons', much like the English expression 'there is only one way to' can be followed by any sort of embedded infinitival. However, as pointed out by the a reviewer, this idiom is attested with other nouns as well, suggesting that *façons* is not an inherent part of the idiom and that *trente-six* may have a special use as an expression indicating an underspecified big number.

(ii) *Je n'ai pas trente-six solutions!*
      I NOT-have NEG thirty-six solutions
      'I can offer only one solution'

(iii) *Pour que les gens tombent amoureux de vous, il n'y a pas trente-six méthodes: il faut faire semblant de s'en foutre complètement.*
'To make people fall in love with you there is just one solution (*lit.* there aren't thirty-six methods): you should pretend that you cannot care less'

(iv) *Quand un garçon comme toi quitte une femme comme moi, il n'y a pas trente-six raisons, il n'y en a qu'une!*
'When a man like you abandons a woman like me, there is just one possible motivation (*lit.* there aren't thirty-six motivations)'

NOT NEG to.touch one hair   of the head of..
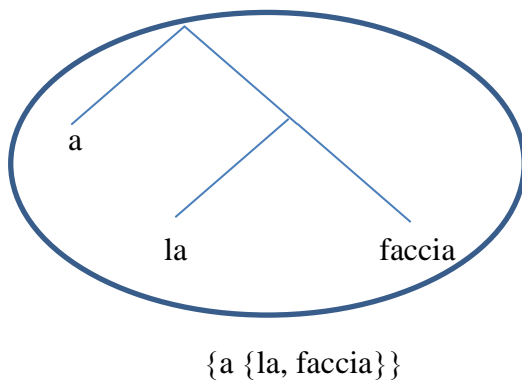'not to harm … in any way'

These are only some examples of PP-less idioms and more will be listed in section 6. But they are representative enough to develop our discussion.

It should be clear that if the nominal domain were parallel to the clausal domain, PP-less idioms would not be expected under Marantz's generalization. Given the solidity of Marantz's generalization in the verbal domain and the independent evidence against the hypothesis that nouns do not take complements (cf. Section 2), we believe that PP-less idioms might be the final nail in the coffin for that hypothesis.

More specifically we propose the derivation in (19) for example in (10): the PP 'del tuo capo' ("of your boss") can be inserted after the definite article + preposition (*alla*) and the noun *faccia* ('face') have merged. The circled area indicates the constituent that gets the idiomatic reading.

(19) Derivation of the idiom: *alla faccia del tuo capo*

Initial Structure

{a {la, faccia}}

Structure after PP merge

{{a {la, faccia}} {del tuo capo}}

We assume that the PP modifier attaches to the unit that is listed in the lexicon as having a special meaning. However, if the noun is not part of an idiom, nothing blocks the PP from attaching to the noun directly. For example, while the idiom *un salto in pizzeria* is given the derivation in (20), the corresponding non-idiomatic expression *tre salti in pizzeria* can be

given the derivation in (21). In (21) the dotted line indicates late merge of the PP modifier. As proposed in previous literature on Late Merge (Lebeaux 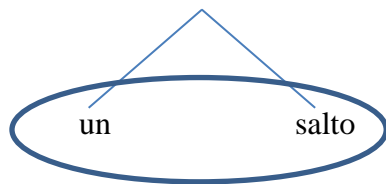1989; Chomsky 1995), while arguments must be inserted cyclically in the derivation because they need to receive a theta-role, adjuncts can be late inserted when the thematic structure is already in place. If the PP that follows the noun does not have the status of an argument (cf. section 2 above), it can be assimilated to an adjunct and be late merged.

(20) Derivation of the idiom: *un salto in pizzeria* ("A brief visit to pizzeria")

Initial Structure

un   salto

{un, salto}

Structure after PP merge

in pizzeria

un   salto

{{ un {salto}} { in pizzeria }}
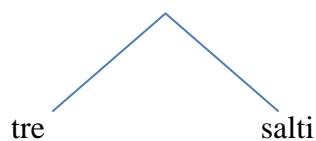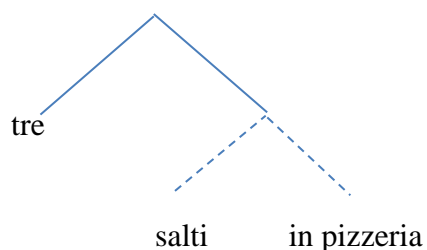
(21) Derivation of the DP: *tre salti in pizzeria* ("Three jumps performed inside a pizzeria")

Initial Structure

tre   salti

{un, salto}

Structure after PP merge

tre

salti   in pizzeria

{ un {salto { in pizzeria }}}

Consistent with the analysis proposed above, PP-less idioms can have an absolute use where the PP is altogether absent:

(22)
SPEAKER 1: *Bill Gates guadagna 500 milioni all'anno*. SPEAKER 2: *Alla faccia!*
SPEAKER 1: Bill Gates earns 500 millions to-the year. SPEAKER 2: To-the face!
SPEAKER 1: 'Bill Gates makes 500 millions a year". SPEAKER 2: 'No less!'

(23) *Sono   uno straccio*
    (I) am   a   rug
    'I feel very bad'

(24) *Farò   un   salto*
    (I) will-make a   jump
    'I will pay a brief visit'

All these cases can be analyzed as cases of idiomatization of D+N without late merged of any PP modifier.

We conclude this section by stressing that the constituent that gets idiomatized in PP-less idioms is a syntactic phrase rather than a complex word.[7] First, while Italian allows N+N compounds, no D+N compound is attested in the language (Bisetto and Scalise 2005), so if PP-less idiom were a compound it would be a very weird one. Second, while we acknowledge that the boundary between compounds and phrases is not a simple issue, the conclusion that PP-less idioms are phrases is supported by a reliable criterion that distinguishes compounds and phrases, namely the impossibility for a noun inside a compound to be modified by an adjective (see Fábregas forthcoming for an articulated defense of this specific criterion). While the *b)* expressions in (25) to (27), in which the adjective modifies a noun belonging to a V+N compound, are ungrammatical, an adjective can modify the noun inside a PP-less idiom, as shown in (28)-(30).

(25) a. *Un apribottiglie*
     A open-bottle PLUR-FEM
     'A bottle opener'
    b. **Un apribottiglie   piccole*
      A open-bottle PLUR-FEM small-PLUR-FEM
      Intended meaning: A bottle opener for small bottles

(26) a. *Un portafogli*
     A   take-paper-PLUR-MASC
     'A wallet'
    b. **Un portafogli   preziosi*

---

[7] There are theories that deny that there is a separation between syntax and morphology. Assuming these approaches, no principled distinction can be drawn between words and phrases. We cannot enter this general debate here, but we want to stress that even these theories must explain why certain categories (words in a traditional sense) obeys the lexical integrity hypothesis, while others (phrases in traditional sense) do not. See Fábregas (forthcoming) for a general discussion of this topic.

A   take-paper-PLUR-MASC   precious- PLUR-MASC
　　　　Intended meaning: A wallet for precious papers

(27)　a. *Una   lavapiatti*
　　　　A   wash-plate-PLUR-MASC
　　A dishwasher
　　b. **Una lavapiatti   incrostati*
　　　　A   wash-plate-PLUR-MASC   crusty- PLUR-MASC
　　　　Intended meaning: A dishwasher for crusty plates

(28)　*Alla   facciaccia brutta di   tuo marito*
　　　To.the   face   ugly of   your   husband
　　　'So much for your ugly husband!'

(29)　*Un salto veloce   a   scuola*
　　　a   skip   quick   to   school
　　　'A quick visit to the school'

(30)　*É   a un piccolo   passo dalla   fama*
　　　Is   at a   small   step from   fame
　　　'(S)he is close to becoming famous'

We therefore conclude that the idiom part in PP-less idioms is phrasal, like *kick the bucket* or *spill the beans*.

5. Other idioms in the nominal domain

Although very simple, the analysis of PP-less idioms that we have proposed in the previous section faces some challenges. They come in two varieties. On the one hand, there are DPs where the N and the PP that follows the noun receive an idiomatic reading while the D is variable. Structurally these idioms are the mirror image of PP-less idioms. On the other hand, there are cases, initially discussed by Bruening (2010), where a verb takes a nominal argument and the idiom is formed by the verb plus the noun while the determiner introducing the noun is not part of the idiom. Finally, there are full DPs that are idiomatized.
We discuss these types in the next three sub-sections.

5.1 The mirror image of PP-less idioms: PP-containing idioms

It is not difficult to find examples where the PP modifier of the noun is part the idiom, while D is not. An Italian example, among many others, is *testa di cuoio* (lit. "head of leather "), which indicates members of special anti-terrorist units. For concreteness, we will call these idioms "PP-containing idioms".

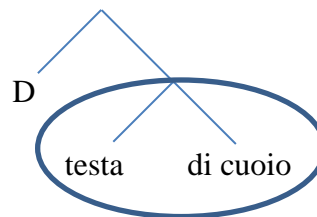(31) *Una   testa   di   cuoio*
　　　A   head   of   leather
　'A member of the anti-terrorist unit'

It is easy to show that the determiner is not part of the idiom, since the idiomatic reading is preserved when the determiner changes.

(32) a. *Nessuna testa   di cuoio*
           No   head   of leather
           'No member of the anti-terrorist unit'
       b. *Molte teste   di cuio*
           Many heads of leather
           'Many members of the anti-terrorist unit'
       c. *Ogni   testa   di   cuoio*
           Every head   of   leather
           'Every member of the anti-terrorist unit'

The problem raised by PP-containing idioms for the analysis proposed in the previous section is that, if the process of idiomatization involves D plus N *before* the PP is merged, then D should be part of the idiom while the PP should not. However, the opposite pattern is observed in PP-containing idioms. A solution would be assuming that the PP *di cuoio* is early merged to the noun *testa* (before the determiner enters the derivation). The circled area indicates the constituent that gets the idiomatic reading under this derivation.

(33)

          {una, {testa {di cuoio }}}

However, the analysis in (33) is not appealing, since it implies two rather different derivational histories in the DP. It amounts to assuming that sometimes the noun is merged to the PP in step one of its derivational history, giving rise to PP-containing idioms, while in other cases it is merged to D, giving rise to PP-less idioms. It is however desirable to assume a uniform syntax for within and derive the existence of the two categories of idioms from some other structure building mechanism.
There is indeed empirical evidence that the right analysis for PP-containing idioms is a different one. The crucial observation is the PP that follows the noun in PP-containing idioms must be formed by a P followed by a *bare* noun, even when the relevant preposition takes full NPs in other (non-idiomatic) contexts.
For example, (34) which is identical to (33) but for the fact that the preposition *di* takes a full NP, is ungrammatical, despite the fact that the same preposition takes a full NP in the non idiomatic context (35).

(34) *Una testa   del   cuoio*
        A   head   of-the leather
(35) *La qualità   del   cuoio*
        The quality   of-the leather
        'The quality of leather'

The following triples show the same point for a variety of PP-containing idioms. In the *a.* example the preposition takes a bare noun and the idiomatic reading emerges. The *b.* example shows that the idiomatic reading is lost (or the structure is utterly uninterpretable) if the preposition takes a full NP. The *c.* example, which is a control, shows that the same preposition can take a full NP in non-idiomatic contexts.

For example, the notorious idiom in (36a), which corresponds to the English idiom *dickhead*, loses its idiomatic meaning when the article is added as in (36b). In fact, (36b), if anything, can only mean the upper part of the penis. The same reasoning applies to the other examples[8].

(36)   a. *testa di cazzo*
          'dickhead'
       b. *#testa del   cazzo*
          head   of-the dick
       c. *la dimensione del   cazzo*
          the dimension of-the dick
          'the dimension of the penis'

(37)   a. *colpo di testa*
          Hit of head
          'impulse'
       b. *#colpo della testa*
          Hit of-the head
       c. una foto   della   testa
          a   picture   of-the head
          'The picture of his head'

(38)   a. *luna   di   miele*
          Moon of honey
          'honeymoon'
       b. *# luna   del   miele*
          moon of-the honey
       c. il sapore   del miele
          the flavor of-the honey
       'The flavour of honey'

(39)   a. *arma   da   fuoco*
          Gun   from fire
          'firearm'
       b. *# arma dal   fuoco*
          Gun   from-the fire

---

[8] In (34), (38), (39) and (40) the PP modifier introduces a material: leather, honey, and straw. A mass noun in Italian can be bare even in non-idiomatic expressions, cf. (i), but it does not need to, cf. (ii).
(i) Il risultato dipende dalla qualità di cuoio usato
The output depends on-the quality of leather used
(ii) Il risultato dipende dalla qualità del cuoio usato
The output depends on-the quality of-the leather used
'The final result depends on the type of leather that has been used'
Furthermore, in other examples, such as (36), (41), (43) and (44), the bare noun is not a mass noun, so it cannot be bare unless it is part of a PP-containing idiom.

c. *il   pericolo del   fuoco*
the   danger   of-the fire
'the danger of fire'

(40)   a. *coda di   paglia*
tail   of   straw
'guilty conscience'
b. # *coda della   paglia*
tail   of-the straw
c. *il colore della   paglia*
the color of-the   straw
'the color of-the straw'

(41)   a. *(fare) orecchie da mercante*
To-make ears from merchant
((to turn) a deaf ear'
b. # *(fare)   orecchie dal   mercante*
To-make ears   from-the merchant
c. *le notizie dal mercante*
the news from-the merchant
'news from the merchant'

(42)   a. *(all')   acqua di rose*
to-the water of roses
'soft'
b. # *(all') acqua delle rose*
to-the water of-the roses
c. il profumo delle rose
the scent of-the roses
'the scent of roses'

(43)   a. *casa da gioco*
house from game
'gaming house'
b. # casa dal gioco
house from-the game
c. la dipendenza dal gioco
the addiction from-the game
'gambling addiction'

(44) a. *colpo di   stato*
blow   of   state
'coup'
b. # *colpo dello   stato*
blow   of-the   state
c. *la vendetta dello   stato*
the revenge of-the state
'the state's revenge'

(45)   a. *frutti di mare*
fruits of sea

'seafood'
   b.# *frutti del   mare*
      fruits of-the sea
   c. le   onde del   mare
      'seawaves'

We can conclude from this observation that PP-containing idioms are impoverished. In particular, in Italian articles and nouns agree in gender and number but prepositions do not agree with the category they select. Therefore, the absence of the article has an important consequence, namely that there is no agreement internal to the PP-containing idiom. In fact, Fábregas (forthcoming) proposed that a key criterion to distinguish words and phrases is that the syntactic operation responsible for agreement can take place inside phrases but not inside (compound) words. This explains among other things the data reported in Section 4, example (27) and repeated below with a new numeration: in languages where nouns and adjectives agree in gender and number, an adjective cannot modify a noun within a compound.

(46)   **Una lavapiatti*   *incrostati*
           A   wash-plate-PLUR-MASC crusty- PLUR-MASC
  Intended meaning: A dishwasher for crusty plates

We propose a similar reasoning for PP-containing idioms: they are complex words and cannot contain an article because the presence of an article would require the syntactic operation of agreement to apply word-internally.

The picture that is emerging is that PP-less idioms are syntactic idioms while PP-containing idioms are morphological idioms, *i.e.* compounds. A reviewer objects to this hypothesis by observing that while PP-containing idioms contain some functional structure, in the form of a preposition, and argues that this prevents their analysis in terms of morphological units. However there are clear instances of (close) compounds in Italian that contain a preposition. Some are listed in (47).

(47) a. *pomodoro*
     apple-of-gold
     'tomato'
   b. *ficodindia*
     figue-of-India
     'prickly pear'
   c. *calcinculo*
     kicks-in-ass
     'chairoplane'
   d. *saltimbocca*
     jumps-in-mouth
     'Italian dish made of veal lined or wrapped with prosciutto and sage'
   e. *saltimbanco*
     jumps-in-desk
     'acrobate'

This non exhaustive list shows that compounds in Italian (unlike in English) can include functional material such as prepositions. Crucially, all these compounds include a bare noun. If we add a determiner into any of the examples in (47), either the structure becomes

ungrammatical or the compound reading is lost. For example, is we add a determiner to *calcinculo*, we obtan *calci nel culo*, a phrase only meaning *kicks in the ass*.

Furthermore, we can offer independent evidence supporting the hypothesis that PP-containing idioms are morphological units as opposed to phrases. We do this by applying other tests that distinguish words from phases. Some of the criteria that distinguish phrases from compounds, like extractability, are not well-suited for the case of idioms, because extractability from idioms varies even in indisputable cases of phrasal categories, with extractability claimed to be totally impossible (*The bucket that you kicked…*) or somewhat marginal (*The beans that were spilled earlier….*).

Another criterion is more useful. Compounds, unlike phrases, are anaphoric islands, namely a constituent internal to the compound cannot be the antecedent of a pronoun (only the entire compound can). This is illustrated by the ungrammaticality of (48a), in which the null subject is intended to refer back to the noun *piatti* (dishes), which contrasts with (48b), a control showing that the same noun can be the antecedent of a pronoun when the noun is part of the verb phrase (the same point holds in English as shown by the deviant sentence "*I bought a dish$_i$washer but it does clean them$_i$/it$_i$").

(48) a. * *Ho   comprato   una [lava-**piatti**],   ma non **li**   lava   bene.*
           Have.1sg bought   a   [wash-dishes],   but not them washes fine
           Intended meaning: 'I have bought a dishwasher, but it doesn't clean dishes properly'

      b. *Ho   comprato [dei   **piatti**], ma **pro** non   mi   piacciono*
          Have.1sg bought   some plates, but pro not   to-me like
         'I bought some plates, but I don't like them'

                                                              (adapted from Fábregas forthcoming)

Crucially, we can show that PP-containing idioms are anaphoric islands while PP-less idioms are not. For example, the noun in the PP-less idiom in (49) can be the antecedent of the pronoun. In the sentences in (49), the pronoun (overt or null) and its antecedent are marked in bold.

(49) a. *Brinderemo   alla   **faccia** della tua capa e   a **quella** del   suo segretario.*
           We-will-toast   to-the face   of-the yourboss and to that   of-the her   secretary
           'We will toast despite your boss and her secretary'

      b. *Eravamo a un **passo** dalla   vittoria e   a   uno **pro** dalla   sconfitta.*
          We-were to a step   from-the victory   and to   one   from-the defeat
          'We were very close to borh victory and deafet'

      c. *Faremo   un **salto** da tua   madre   e   uno **pro** da mio padre*
          We-will-make a   jump   to your   mother   and one   to   my father
          'We will pay a short visit to your mother and to my father'

Admittedly, other cases of anaphora are marginal, but still acceptable to our judgment:

(50) ? *Alla   **faccia** del   capo, che poi **pro** è   anche brutta*
         To-the face   of-the bossc that after   is   also   ugly

We conjecture that the oddity of sentences like (50) is not to be attributed to the anaphoric islandhood of PP-less idioms but to a lack of semantic parallelism, since the same noun is used in the same sentence first idiomatically and then literally (through pronominalization).

Be that as it may, PP-containing idioms are always strong anaphoric islands, in sharp contrast with PP-less idioms. All the sentences in (51) are gibberish.

(51) a.# *Siamo andati in **luna** di miele, ma non **l'**abbiamo vista perché c'erano molte nuvole.*
We-have gone to moon of honey but not it-have seen because there were many clouds
Intended meaning: We went to honeymoon but we did not see it (=the moon) because it was cloudy
b.* *Lei sogna la **luna** di miele, io **quella** di Giove.*
She dreams the moon of honey I that of Jupiter
Intended meaning: she is dreaming her honeymoon and I am dreaming that (=moon) of Jupiter
c. #*Ho incontrato un vero **cuor** di leone anche se personalmente **gli** preferisco il fegato.*
I-have met a true heart of lion even if personally to-him I-prefer the liver
Intended meaning: I met a very brave person (lit. heart of lion) although I prefer liver

The following contrast is a revealing minimal pair. The sentences in (52) include a PP-containing idiom with the characteristic impoverished structure discussed above (the preposition *di* takes the bare noun *mare*). The expression in (53) is identical but for the fact that the preposition takes a full PP (*del mare*). The DP in (53) refers to all products of the sea, while the PP-containing idiom in (52) is the canonical translation of the English word *seafood*. Anaphora is blocked in (52) but possible in (53). So, as expected, anaphoric islandhood goes hand to hand with lack of article, confirming that PP-containing idioms are indeed compounds.[9]

(52) a. #*Mangio molti frutti di **mare**, ma non **ne** mangio le alghe*
I-eat many fruits of sea but not of-it I-eat the weeds
Intended meaning: I eat seafood (lit. *fruits of the sea*) but I do not eat seaweed (lit. its weed)
b.#*Mi piacciono molto i **frutti** di mare, ma non **quelli** dell'albero.*
To-me like a-lot the fruits of sea but not those of-the tree
Intended meaning: I like seafood (lit. *fruits of the sea*) a lot but not those (=fruits) by the trees

(53) a. *Mangio molti frutti del **mare**, ma non **ne** mangio le alghe*
I-eat many fruits of-the sea but not of-it I-eat the weeds
'I eat seafood but I do not eat seaweed'
b. *Mi piacciono molto i **frutti** del mare, ma non **quelli** dell'albero*
To-me like a-lot the fruits of-the sea but not those of-the tree
'I like seafood a lot but not fruits of trees'

In this section, we have discussed the mirror image of PP-less idioms. In PP-less idioms D+N receive an idiomatic reading while the PP following the noun can vary. In PP-containing idioms, N+PP conveys an idiomatic reading while D can vary. We have seen clear evidence that idiomatization takes place in the morphological component for PP-containing idioms but

---

[9] A reviewer wonders whether the contrast could not be due to the fact that bare nouns in general cannot entertain anaphoric relations. This does not seem to be the case, as the example in (i) clearly shows.
(i) Ho visto **ragazzi** studiare fino all'alba, ma **pro** erano pochi.
Have seen guys study till dawn, but pro were few
'I saw people studying 'till dawn, but they were very few'.

it takes place in the syntactic component for PP-less idioms. All in all, PP-containing idioms do not require a complication to the approach proposed to PP-less idioms in Section 4.

5.2 Full DPs idioms

In addition to PP-less idioms and PP-containing idioms, a third type of idiom in the nominal domain is allowed under the approach that we have been pursuing. These are idioms in which the entire nominal constituent (D+N+PP) gets the idiomatic reading. Since the entire DP is a constituent, nothing prevents the process of idiomatization from applying to it. Idioms of this type do exist. Two examples are *la tana del lupo* (*lit*. the den of the wolf), meaning "lion's den" and *ogni morte di papa* (lit. every death of pope), meaning "very rarely". In these idioms the PP and D are both also fixed (cf. 55a, which only gets a literal reading and 55b , which is unacceptable):

(54) a. *la   tana   del   lupo*
          The-SING den-SING   of-the wolf
          'lion's den'
      b. *ogni morte di papa*[10]
          every death of pope
          'very rarely'

(55) a. *#una   tana del   lupo*
             A   den of-the wolf
          'One of the wolf's dens'
      b. *la morte   di papa
             the death   of pope

To summarize, there are three ways in which a (part of) a nominal constituent can become an idiom:
(i) idiomatization involves the syntactic constituent D + N before the PP is merged (PP-less idioms)
(ii) idiomatization involves the morphological category N + P + bare noun (PP-containing idioms)
(iii) idiomatization involves the entire syntactic constituent DP

---

[10] In this idiom P must take a bare noun, as in PP-containing idioms. This is shown by the oddity of (i), which presupposes (counterfactually) that the same pope can die several times.
(i) *Succede   ogni morte del   papa*
    it-happens   each death   of-the   pope
"It happens each time the (present) pope dies"
This suggest that *ogni morte di papa* might be a compound, as other PP-containing idioms. This analysis would be consistent with the observation that it is an anaphoric island:
(iii) # *Succede   ogni morte di **papa**   ma per il   momento **pro**$_i$ sta bene*
        it-happens   each death of pope   but for the   moment   is healthy
        Intended meaning: 'It happens rarely but for the moment the pope is healthy'
Why should a morphological idiom contain the determiner (*ogni*) in this specific case? In fact, *ogni* has a fixed default form (it is not inflected for number and gender, unlike articles and other determiners). Therefore, no syntactic operation of agreement is required when *ogni* is merged with *morte di papa*. For this reason, we conjecture, idiomatization can be morphological in this specific case although it targets a full D.

If we are on the right track, another type of idiom should *not* be attested, namely idioms formed by N and PP where the PP that modifies the noun takes a full NP. However, although much rarer than the PP-containing idioms we discussed in Section 5.1, they seem to be attested. We could find two candidates, namely *spina nel fianco* ("thorn in my side") and *avvocato del diavolo* ("devil's advocate"):

(56) a. *La/una/ogni   spina nel   fianco*
          The/a/every   thorn in-the   side
      b. *L'/un/ogni   avvocato del   diavolo*
          The/a/every   advocate of-the   devil

To reiterate, these cases are a challenge to our approach because D is variable (so they cannot derived as in *i)* or in *iii)*) and P does takes a bare noun (so they cannot be derived as in *ii)*). However, these idioms are special in another respect. The noun (*spina* and *avvocato*) can be plural and the idiomatic reading is still allowed:

(57) a. *Obama e   Michelle   sono   le   spine   nel   fianco di Trump*
          Obama and   Michelle   are   the thorns in-the side   of Trump
          'Obama and   Michelle are a nuisance for Trump'
      b. *Non abbiamo bisogno di avvocati   del   diavolo in   questa azienda!*
          Not we-have need   of lawyers   of-the devil   in   this   company
          'In this company an devil's advocate is not needed!'

We double checked our own intuition with a Google search and the results confirm it, since for both idioms about 10% of Google hits contain a plural noun (*spine* and *avvocati*), too big a percentage to be considered noise.[11] Of course, in uncontroversial cases of idioms, the variability of the noun in number is not found (cf. #*kick the buckets*) and this extends to the PP-less idioms we previously considered.

(58) a. # *Gianni e Maria   sono due stracci*
            Gianni and Maria are   two rugs
      b. #*Alle   facce della   tua capa   e   del   suo segretario*
            To-the faces of-the   your boss   and of-the her   secretary
      c. # *Faccio   due salti   in   farmacia   e   in   libreria*
            I-make   two jumps   in   pharmacy and   in   library

Based on this pattern, we conclude that *spina nel fianco* ("thorn in my side") and *avvocato del diavolo* ("devil's advocate") are not fully idiomatized expressions. This explains their unexpected variability in the choice of the determiner.

---

[11] Search on the Google (Italian Version) performed on July, 24 2018:
spina nel fianco 168000 hits
spine nel fianco 14200 hits
avvocato del diavolo 164.000 hits
avvocati del diavolo 16.900 hits
Notice that these expressions cannot be productively used with a compositional meaning. A random control shows that the overwhelming majority of both singular and plural uses are given the idiomatic reading.

5.3 Idioms that do not look like constituents

Our approach to idiom formation has been fairly conservative in one respect. We have been assuming that words (including complex words that behave like compounds) and constituents created by syntax can acquire a meaning (related or unrelated to their compositional meaning) that must be listed in the mental lexicon. What is more innovative in our approach is the assumption that D and N merge in the first step of the derivation and any PP that modifies the noun is merged later, being assimilated to an adjunct.

However, the idea that what can get idiomatized is a constituent has been challenged, notably by Bruening (2010) and O'Grady (1999). The examples that Bruening exploit in his paper against the constituency theory of idiom formation are very interesting and, being directly relevant for our discussion of PP-less idioms, they need to be discussed here. A representative case is (59). The variability of the determiner shows that the idiom in (598) is 'pull strings'. This is hard to explain under the cyclic theory of idiom formation because the determiner is part of the constituent that should receive the idiomatic meaning.

(59) a. *Pull **some** strings*
   b. *Pull **a few** strings*
   c. *Pull **yet more** strings*

In order to account for cases like this, Bruening develops a theory of idiom formation that exploits the notion of selection rather than that of constituency. Roughly speaking his idea is that what can become an idiom is the category formed by a selector and its selectee (and the selectees of the selectee, if present). If the verb ('pull' in the case at hand) selects noun phrases, not determiners (*pace* the DP hypothesis), the theory that exploits the notion of selection can account for the pattern in (59).

However, idioms like (60) are a challenge for Bruening's theory:

(60)   *Strike while the iron is hot*

In (60), under standard assumptions, there is no selection relationship between the adverbial phrase 'while the iron is hot' and the verb 'strike', and yet they do form an idiom. Bruening handles this problem by stipulating that the adverbial phrase selects the verb 'strike' but this requires stretching the notion of selection in a way which does not seem desirable.

Can the examples in (59) be reconciled with the approach that we have been developing? A natural possibility is maintaining with Bruening that the idiom is 'pull strings' but denying the D status to the (alleged) determiner of (59). This is suggested by the fact that the (alleged) determiners present in (59) have a clear adjectival use in sentences like (61).

(61)   a. The *attendees were some (but not many)*
      b. *The attendees were a few*
      c. *The attendees were yet more*

If we are dealing with adjectives and not determiners, the cases in (59) admit an analysis compatible with the approach that we have been developing. In particular, the adjective (being an adjunct) can be late merged to the V+N category that gets idiomatized.

If we are on the right track, the real test case to decide between the cyclic theory of idiom formation and the alternative theory proposed by Bruering would be idioms similar to (59)

but for the fact that there is an unambiguous determiner, namely one that never admits an adjectival use. The idioms in (62) and (63) are two candidates, since the definite determiner never allows an adjectival use.

(62)   *Flip the bird*
(63)   *Hit the hay*

Interestingly, this type of idioms the determiner is fixed, as predicted by the constituency theory of idiom formation:

(64)   a. *#Flip a bird*
        b. *#Flip many birds*
        c. *#Flip no bird*

(65)   a. *#Hit some hay*
        b. *#Hit much hay*
        c. *#Hit no hay*

We conclude that the case against the cyclic theory of idiom formation raised by Bruening (2010) is not conclusive.

6. Discarding some alternative analyses for PP-less idioms

Some PP-less idioms appear to be compatible with an alternative analysis where the idiomatized PP-less idiom is a constituent for independent reasons and does not force a late adjunction analysis of the PP. In this section we discuss in particular two of these analyses: as measure phrases and as inverted predicates.

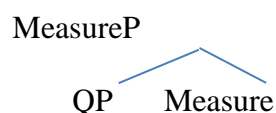### 6.1 Distinguishing PP-less idioms and measure phrases

There are some PP-less idioms that look like measure phrase (Doetjes & Rooryck 2003).
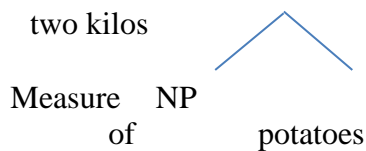
An example is (66).

(66)   *Quattro gatti di spettatori*
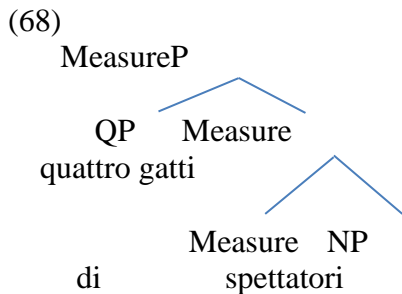   Four   cats of spectators
        'A very small audience'

For this idiom we might extend the measure analysis for uncontroversial cases of measure phrases (cf. Corver 2006), which is illustrated in (67).

(67)

    MeasureP
       /\
     QP   Measure

```
     two kilos
                    /\
   Measure    NP
        of          potatoes
```

If we extend this analysis to the idiom in (66), we obtain the representation in (68).

(68)
```
     MeasureP
              /\
    QP    Measure
  quattro gatti
                 /\
          Measure   NP
     di          spettatori
```

If we adopt the representation in (68), the element which gets the idiomatic reading ('quattro gatti') is a QP in the specifier of the measure phrase, therefore it is an independent constituent.
One can ask whether the same analysis can be extended to the PP-less idioms like (15), repeated as (69). We can exclude this. First, the PP that follows the noun does not express in any possible sense a substance that can quantified, since the PP has a locative, instrumental or temporal meaning:

(69)   *Un salto (in città,   al   mare,   in vespa, stasera ...)*
       A skip   (in town,   to-the sea,   in vespa, tonight)
       'A brief visit downtown, to the seaside, by using vespa, tonight'

Second, measure QPs like 'due chili' ("two kilos") can be predicated of the noun they quantify: as shown in (70), a measure QP like 'due chili' can appear in the canonical position of the predicate.

(70)   *Le   patate   sono due   chili*
       the potatoes   are two   kilos
       'there are 2 kilos of potato'

This is also true of idioms like 'quattro gatti', for example: in (71) 'quattro gatti' is predicated of the noun it quantifies, sitting in a canonical postcopular predicate position.

(71)   *Gli spettatori   sono   quattro   gatti*
       the spectators   are   four   cats
       'The audience is very small'

However, with PP-less idioms like *un salto* in (72a) it is totally impossible for the idiomatic phrase to sit in the canonical predicate position (cf. 72b). This appears to be a confirmation that the measure phrase analysis is not a viable option with these PP-less idioms.

(72)   a. *Un salto   in città*
          a   jump   in town
          'A brief visit to town'

b. *In città   è un salto*
   In town   is a jump

In this section we have argued that, although some idioms may be analysed as belonging to measure phrases, other PP-less idioms resist this analysis. In the next section we consider another possible analysis for PP-less idioms, which is related but distinct from the measure phrase analysis. This is predicate inversion analysis.
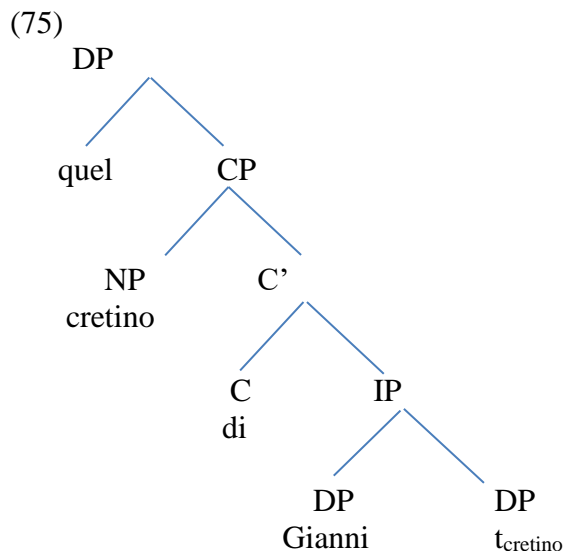

6.2 Distinguishing PP-less idioms and inverted predicates

In Italian and other Romance varieties a very productive construction involves so-called predicate inversion: in addition to a canonical subject-predicate construction like (73) a DP like (74) can be found, in which the predicate 'cretino' ("stupid") is promoted to a position before the subject of predication 'Gianni'.

(73)   *Gianni è un cretino*
        Gianni is a stupid
        'Gianni is stupid'

(74)   *Quel cretino di Gianni*
        That stupid   of Gianni
        'How stupid Gianni is'

These DPs are fully natural only if introduced by a demonstrative. These constructions have been the object of a fairly extensive literature, which stems from an influential analysis proposed by Kayne (1994). For concreteness we give a schematic representation of Kayne's analysis in (75).
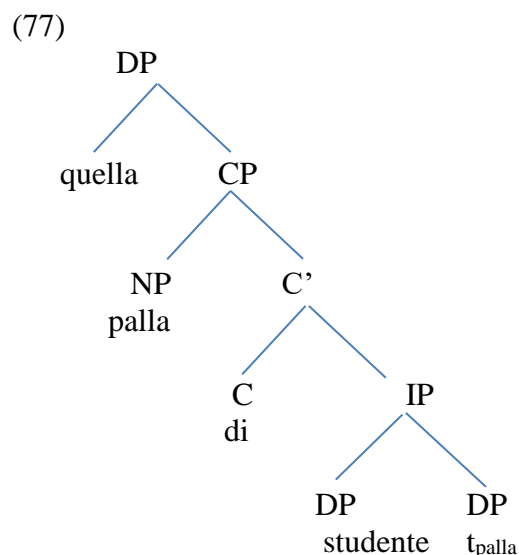
(75)

In (75) the predicate 'cretino' is generated in its canonical position in which it is in a local configuration with the subject of predication and later moves to Spec, CP.
The reason why we mention predicate inversion in the context of our discussion of idioms is that the inverted predicate analysis *prima facie* can be extended to some PP-less idioms. We

show this by taking as a representative the idiom in (76) which might receive the predicate inversion analysis schematically represented in (77).

(76) *Quella palla   di studente/professore/cliente*
     that   ball   of student/professor/customer
        'A very boring student/professor/customer'

(77)
```
                DP
              /    \
         quella    CP
                  /  \
                NP     C'
               palla  /  \
                     C     IP
                     di   /  \
                        DP     DP
                    studente  t_palla
```

In (77) the noun that gets an idiomatic reading ('palla') is analysed as the underlying predicate of the DP 'studente' and moves leaving a trace in its base position, much like the predicate does in (75).
If the analysis in (77) is correct, the idiom is identified with the inverted predicate 'palla', which sits in Spec,CP.
However, not all PP-less idioms can be analyzed as involving inverse predication. There are cases where the PP that is excluded from the idiomatic reading cannot appear in the canonical subject position of the predicate. For example, a PP-less idiom in which the predicate inversion analysis is precluded is 'due passi da…', namely "few steps from…", in (78): (78) shows that its alleged input is ungrammatical.

(78)   *Due passi da casa*
        two steps from home
        'very close to home'
(79) **Casa   è due passi*
        home is two steps

Another clear piece of evidence showing that the predicate inversion analysis cannot be extended to all PP-less idioms comes from the fact that the PP is not obligatory, as indicated in the examples in (80) to (84).

(80)   *Facciamo   un salto*
        Let-us-make a   jump
        'Let us pay a short visit'
(81)   *Farò   due passi*
        I-will-make   two steps

'I will take a walk'

(82)   *C'erano   si e no   quattro gatti*
       There-were   yes and no four   cats
       'There was almost no one'

(83)   *Gianni è una palla*
       Gianni is a ball
       'Gianni is very boring'

(84)   *Suo marito   è   uno straccio*
       Her husband is   a   rug
       'Her husband is in a bad shape'

The grammaticality of (80) to (84) cannot be explained under the predicate inversion analysis for a very simple reason, which becomes apparent if one goes back to the representation in (77): the subject of predication (*studente* in 77) is internal to the PP. Since in (80) to (84) there is no PP, there is no subject of predication and a fortiori no predication analysis is viable for these examples.

We have now the necessary background to deal with an interesting fact, namely the fact that with some idioms, including the idiomatic reading of *palla* ("ball", i.e. "something boring"), a limited range of determiner variability is admitted: the indefinite determiner *uno/una* can alternate with the demonstrative *quello/quella*, and both (85) and (86) are fully acceptable. However, other determiners are totally out in the idiomatic reading (cf. 87 to 89).

(85)   *(Ho   incontrato) una palla di   studente*
       (I) have   met   a   ball of   student
       'I met a very boring student'

(86)   *(Ho   incontrato) quella palla   di   studente*
       (I) have met   that ball   of   student
       'I met the student, who is very boring'

(87)   # *(Ho   incontrato) la   palla   di studente*
          (I) have   met   the   ball   of student

(88)   # *(Ho   incontrato) alcune palle di   studenti*
          (I) have   met   some   balls of   students

(89)   # *(Ho   incontrato) molte palle di studenti*
          (I) have   met   many balls of students

Prima facie, the pattern in (85) to (89) is puzzling. Since there is some determiner variability, one would be inclined to conclude that the idiomatic category is just the noun *palla*, not the D+N category. On the other hand, if one assumes this, it remains mysterious why the idiomatic reading is not possible with other determiners. However, we think that this pattern can be fully understood after our discussion of predicate inversion. While (85) is a PP-less idiom derived as proposed in Section 4 (namely merge of the PP *di studente* to the category D+N that gets idiomatized), (86) is a case of predicate inversion, as suggested by the fact that it is introduced by a demonstrative, namely the determiner that most naturally occurs with an inverted predicate.

This analysis is strongly supported by the observation that if the PP that follows the idiomatic noun is omitted, the idiomatic reading is no longer allowed with a demonstrative determiner

(cf. 90). This is expected if the idiomatic reading of *quella palla* indeed results from predicate inversion, because in the absence of the PP no subject of predication is available, as we explained above. On the other hand, as we already noticed, the idiomatic reading is retained in (91) where the idiom is introduced by the indefinite article. This is expected if (91) is a D+N idiom to which a PP can be late-merged if present.

(90) #*Gianni è   quella palla*
         Gianni is   that   ball
(91) *Gianni è   una palla*
        Gianni is   a   ball
        'Gianni is very boring'

A similar pattern can be found with another idiom (*straccio*, "rug") that allows the indefinite determiner/ demonstrative alternation when a PP is present (cf. 92). Without the PP, the idiomatic reading is possible with the indefinite determiner (cf. 93) but impossible with the demonstrative (cf. 94):

(92) *Ho   incontrato uno/quello straccio d'uomo*
        I-have   met   a/that   rug   of man
        'I met a/the poor guy'
(93) *Gianni è   uno   straccio*
        Gianni is   a   rug
        'Gianni is in a very bad shape'
(94) #*Gianni é   quello straccio*
        Gianni is   that   rug

We can summarize this section by saying that some idioms can result from predicate inversion. However, not all PP-less idioms can be analyzed in this way. More specifically, PP-less idioms introduced by the indefinite determiner are not inverted structures, as shown by the fact the PP is optional. The PP is obligatory with PP-less idioms introduced by a demonstrative because these idioms are inverted predicate structures and the PP contains the subject of predication.

7. Conclusion

Our investigation of nominal idioms has been prompted by a research agenda that focuses on the derivational history of DPs. Based on independent evidence, we have been exploring the unorthodox hypothesis that the first step of the derivation involves merge of D and N, while any modifier of N (including PPs that are treated as complements of the noun) enters the derivation in later steps.
We started this paper by defending the well-known generalization that the verb and its object can form an idiom to the exclusion of the subject while the verb and its subject cannot form an idiom to the exclusion of the object. This generalization supports the hypothesis that only constituents can become idioms.
If in principle any constituent can become an idiom and if the first step of the derivation involves merge of D and N, idioms should exist that reflect the constituent created at this step, namely structures where D and N receive an idiomatic reading while any PP that modifies the noun does not. Taking Italian as a case-study, our main empirical contribution in this paper is showing that such idioms (PP-less idioms in the terminology that we have been using) do exist.

After identifying PP-less idioms, we switched to their mirror image, namely cases where N+PP are part of the idiom while D is variable. We were able to show that these expressions (that we labeled PP-containing idioms) are the output of the morphological component, a desirable conclusion, since it allows us to maintain a uniform theory of structure formation in the syntax of the nominal domain.

In the last part of the paper we considered cases of idioms that are recalcitrant to the analysis that we have been pursuing but we could show that they do not affect the main conclusion that we want to draw, namely that there are idioms that contribute to the evidence that the verbal and the nominal domain differ, since verbs take complements, while nouns do not.

We are well aware that this paper has a serious limitation, namely it mainly focuses on a single language (Italian). However, as we have seen, idioms are a complex object of research, so an in-depth investigation was needed, and this took the entire space of the paper. More importantly, it is well known that one tends to find what he or she is looking for. As PP-less idioms are unexpected if one assumes a parallelism between clausal and nominal domain, they were not looked for. We are confident that once this new species of idioms has been identified in a language it will be easier to chase it in other languages. More PP-less idioms are yet to come.

References

Abney, Stephen P. 1987. The English noun phrase in its sentential aspect. Ph.D. Dissertation, MIT.

Adger , David. 2013. *A syntax of substance*. Cambridge, MA: MIT Press.

Baker, Mark. 2009. The macroparameter in a microparametric world. In T. Biberauer (ed.) *The limits of language* variation, 351-364. Amsterdam : Benjamins.

Belletti, Adriana and Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6: 291-352.

Bernstein, Judy B. 2008. The DP Hypothesis: Identifying Clausal Properties in the Nominal Domain. In Mark Baltin & Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, Blackwell Publishers Ltd.

Bisetto, Antonella and Sergio Scalise. 2005. The classification of compounds. *Lingue e Linguaggio*, IV(2), 319–332.

Bruening, Benjamin. 2010. Ditransitive asymmetries and a theory of idiom formation, *Linguistic Inquiry* 41: 519-562.

Cecchetto, Carlo and Caterina Donati. 2015. *(Re)labeling*. Cambridge MA: MIT Press.

Chomsky, Noam. 1970. Remarks on nominalization. In: Jacobs, Roderick & Rosenbaum, P. (eds.) *Reading in English Transformational Grammar*, 184-221. Waltham, MA

Chomsky, Noam. 1995. *The Minimalist Program*, Cambridge, MA: MIT Press.

Chtareva, Angelina. 2004. Experiencer Analysis of Subject Idioms in Russian. In Steven Franks, Frank Y. Gladney, Mila Tasseva-Kurktchieva_ (eds.) *FASL 13 Proceedings*, 80-92. Ann Arbor: Michigan Slavic Publications.

Corver, Norbert. 2006. Getting the syntactic measure of measure phrases. *The Linguistic Review* 26: 67-134.

Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the Definition of Word*, Cambridge (Mass.), MIT Press.

Doetjes Jenny & Johan Rooryck. 2003. Generalizing over Qualitative and Quantitative Constructions. In Martine Coene & Yves D'Hulst (eds.) *From NP to DP*: 277-295. Amsterdam: John Benjamins

Dowty, David. 2003. Thematic proto-roles and argument selection. In Javier Gutiérrez-Rexach (ed.) *Semantics: critical concepts*, IV: 221-244. London: Routledge.

Fábregas Antonio. forthcoming. On a grammatically relevant definition of word and why it belongs to syntax. Available at: https://uit.academia.edu/AntonioF%C3%A1bregas http://www.academia.edu/667065/On_a_grammatically_relevant_definition_of_word_and_why_it_belongs_to_syntax (assessed July 2018)

Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.

Hale, Ken and James Keyser 2002. *Prolegomenon to a theory of argument structure*. Cambridge MA: MIT Pres.

Harley, Heidi and Megan Stone. 2013. The 'No Agent Idioms' hypothesis. In Raffaella Folli, Christina Sevdali, and Robert Truswell (eds.) *Syntax and its limits*, 251-275. Oxford: Oxford University Press.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16: 547-593.

Horvath, Julia & Tami Siloni. 2009. Idioms: Mental Representation and Acquisition. Unpublished manuscript, Tel Aviv University. Unpublished ms., Tel Aviv University.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge MA: MIT Press.

Kayne, Richard. 2009. Antisymmetry and the lexicon. *Linguistic Variation Yearbook* 2008: 1-32.

Lebeaux, David. 1989. *Language Acquisition and the Form of Grammar*. Ph.D. Dissertation, University of Massachusetts, Amherst.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge MA: MIT Press.

Wasow, Thomas, Ivan A. Sag, and Geoffrey Nunberg. 1984. Idioms: An interim report. In *Shiro Hattori and Kazuko Inoue (eds.) Proceedings of the XIIIth International Congress of Linguistics*, 87-96. Tokyo: Nippon Toshi Center.

O'Grady, William. 1998. The syntax of Idioms. *Natural Language and Linguistic Theory* 16: 279-312.

Pesetsky, David. 1995. Zero Syntax. Experiencers and cascades. Cambridge MA: MIT Press.

Williams, Edwin. 2007. Dumping lexicalism. In Gilliam Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, Oxford: OUP, 353–382.