

# Quattro passi into the DP structure

*Carlo Cecchetto*

University of Milan-Bicocca

[carlo.cecchetto@unimib.it](mailto:carlo.cecchetto@unimib.it)

*Caterina Donati*

Université Paris Diderot

[cdonati@linguist.univ-paris-diderot.fr](mailto:cdonati@linguist.univ-paris-diderot.fr)

# TWO MAIN SYNTACTIC THEORIES OF IDIOM FORMATION, I

Oversimplifying the picture, there are two main syntactic theories of idiom formation.

## *I. Constituency Theory*

The most traditional account for possible and impossible idioms in terms of constituency: an idiom must form a constituent (in order to be able to be stored as a chunk in the lexicon).

# TWO MAIN SYNTACTIC THEORIES OF IDIOM FORMATION, II

## II. *Selection Theory* (Bruening 2010, but also O'Grady 1999)

### **The Principle of Idiomatic Interpretation**

X and Y may be interpreted idiomatically only if X selects Y.

### **Constraint on Idiomatic Interpretation**

If X selects a lexical category Y, and X and Y are interpreted idiomatically, all of the selected arguments of Y must be interpreted as part of the idiom that includes X and Y (lexical categories are V, N, A, and Adv).

## WHEN BOTH THEORIES AGREE

Both theories correctly capture the fact that no idiom exists that includes the subject and verb but not the object. The constituency theory does that because subject and verb do not form a constituent that excludes the direct object.

The selection theory captures the same fact since any idiom that includes the subject and the verb will have to include all of the selected arguments of the verb, as well.

## WHERE THEY DIVERGE, I

Bruening claims that the selection theory of idioms is superior to the constituency theory of idioms because of examples like (1), in which the idiom ('pull strings') does not include the determiner. This follows if the verb selects noun phrases, not determiners (pace the DP hypothesis).

(1)

Pull *some* strings

pull *a few* strings

pull *yet more* strings

## WHERE THEY DIVERGE, II

The variability of the determiner in idioms like (1) is hard to explain under the constituency theory of idioms because the determiner which varies is part of the constituent that should receive the idiomatic meaning.

(1)

Pull *some* strings

pull *a few* strings

pull *yet more* strings

# A POTENTIAL PROBLEM FOR THE SELECTION THEORY

The selection theory of idioms has to make unorthodox assumptions about selection for the following cases:

- (2) To be in hot water.
- (3) Strike while the iron is hot

In (2) the modifier 'hot' *is* part of the idiom although, under standard assumption, it does not select the noun it forms an idiom with. Similarly, in (3) under standard assumption there is no selection relationship between the adverbial phrase *while the iron is hot* and the verb *strike* but they do form an idiom.

## UNORTHODOX ASSUMPTIONS

Bruening (2010) handles this problem by stipulating that a modifier selects its modiffee. For example, the adverbial phrase *while the iron is hot* would select the verb *strike* in (3).

(3) Strike while the iron is hot

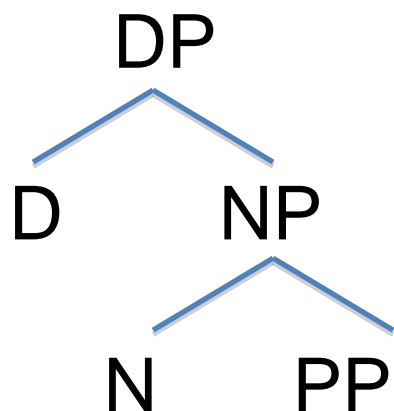


## AIM OF THIS TALK

Aim of this talk is to discuss some idioms that both the constituency and the selection theory of idioms *prima facie* predict to be impossible. We will show that what is wrong is not the theory/theories of idiom formation in itself but the structure of nominal expressions that is standardly assumed.

# THE PROBLEM

## IF THE DP HYPOTHESIS IS RIGHT

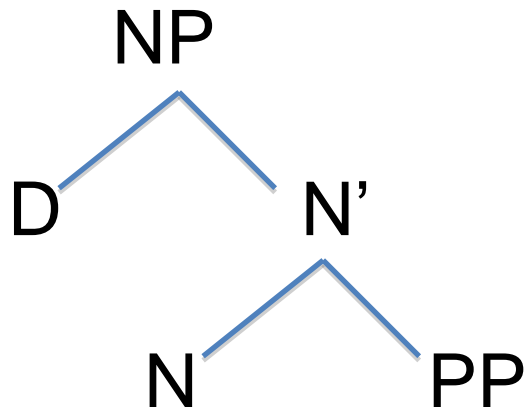


*Theory I. (Constituency):* if D and N are part of an idiom, then the complement PP needs to be part of the idiom as well (the minimal constituent containing both D and N is DP, and it contains the PP).

*Theory II. (Selection):* if D and N are part of an idiom, then the complement PP needs to be part of the idiom as well (D selects N, hence all the arguments selected by N must enter the idiom).

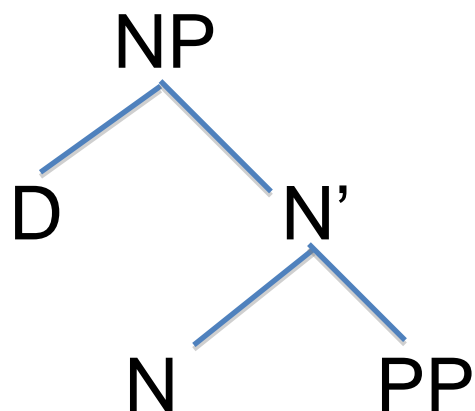
## IF THE DP HYPOTHESIS IS WRONG, I

In a pre-Abney's picture, D is not the head of the nominal expression, but rather its specifier.



*Theory 1. (Constituency):* same prediction. The only constituent containing both D and N also contains PP. Therefore the complement of the noun needs to be part of the idiom.

## IF THE DP HYPOTHESIS IS WRONG, II



*Theory II. (Selection):* D is not a head but a specifier. However, in Bruening's theory two categories may form an idiom only if they are in some selection configuration. Suppose that D selects N (selection and projection are disjoint in Bruening's theory). By the Constraint on Idiomatic Interpretation, the argument of N must be interpreted as part of the idiom that includes D and N  $\rightarrow$  same prediction.

## IDIOMS THAT SHOULD NOT EXIST....

No matter what theory of idiom formation and of NP/DP structure is assumed, it seems that no idiom is predicted to exist of the form “D N PP” in which D and N are fixed while the “complement” of N (or, more generally, the PP selected by N) is variable.

## ... BUT EXIST

This prediction is wrong, since there are indeed idioms of the form “D N PP” in which D and N are fixed while the “complement” of N (or, more generally, the PP selected by N) is variable.

(4) *du fili* (*di spaghetti, di carbonara, di matriciana, al pesto, di insalata, di X* (food with string-like shape))

*Lit.* Two strings of...

‘A portion of X’

In (4) the D is fixed. If ‘*du*’ (two) is replaced by ‘*tre*’ (three) the idiomatic meaning (“a portion of”) is lost. The noun ‘*fili*’ is also part of the idiom, because it loses its original meaning (‘string’).

## OTHER PROBLEMATIC CASES

This example is not isolated. Similar cases include *quattro gatti* (four cats, i.e. very few), or *un salto* (a skip, i.e. a brief visit).

(5) a. Un salto (in città, al mare, in vespa, a/in X)

A skip (in town, to the sea, in vespa, in/to X)

*A brief visit to X or by X*

b. Quattro gatti (di spettatori, di manifestanti, di studenti, di X)

Four cats (of spectators, of protesters, of students)

*Very few X*

Here again the D or Num is fixed, and the noun is also part of the idiom, having most its original meaning.

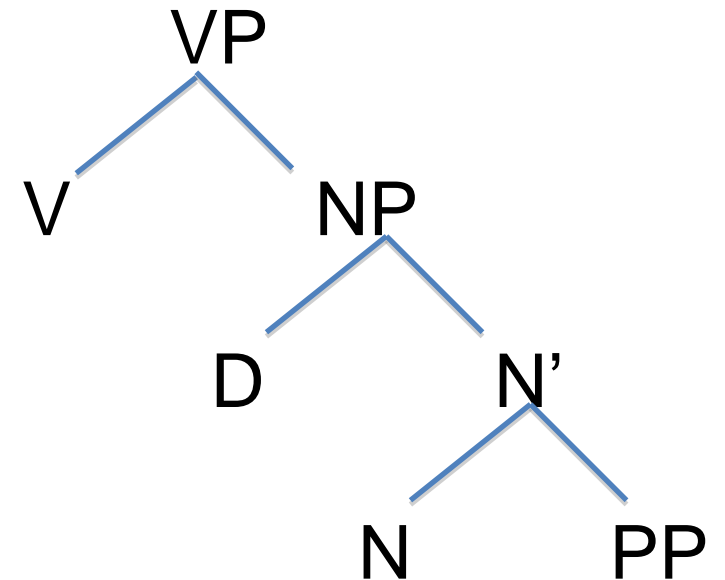
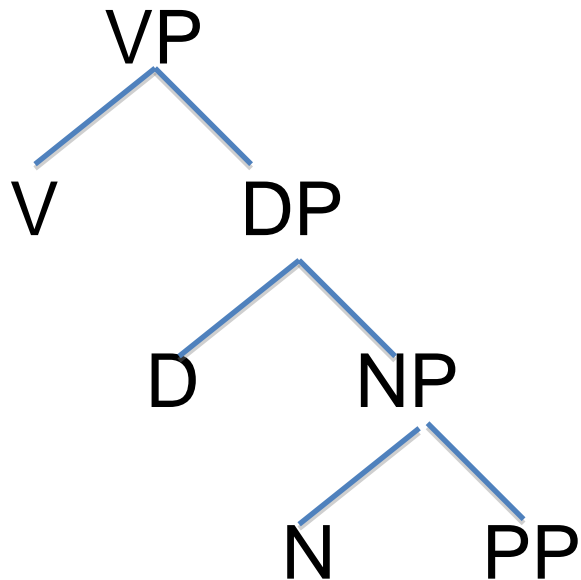


## SUMMARY SO FAR

The conclusion is that no matter whether the DP hypothesis is correct or not, neither the constituency theory of idioms nor the selection theory of idioms, can account for the existence of idioms where D and N form an idiom that excludes the PP selected by N. So, there is a problem.

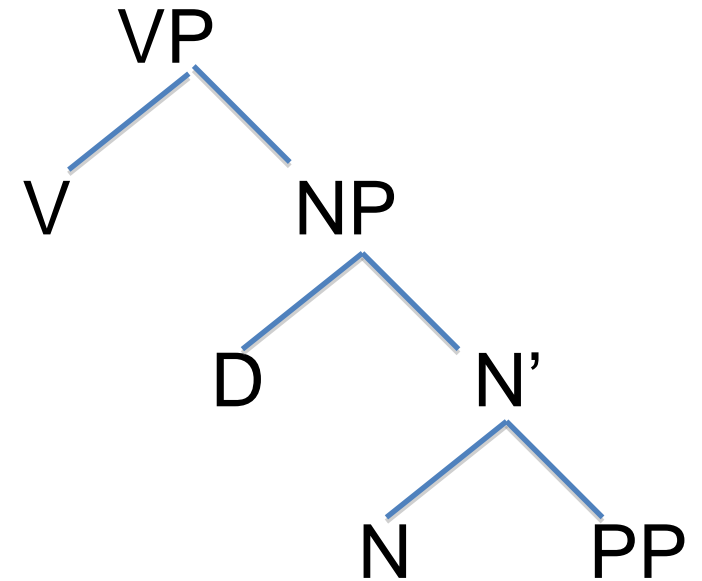
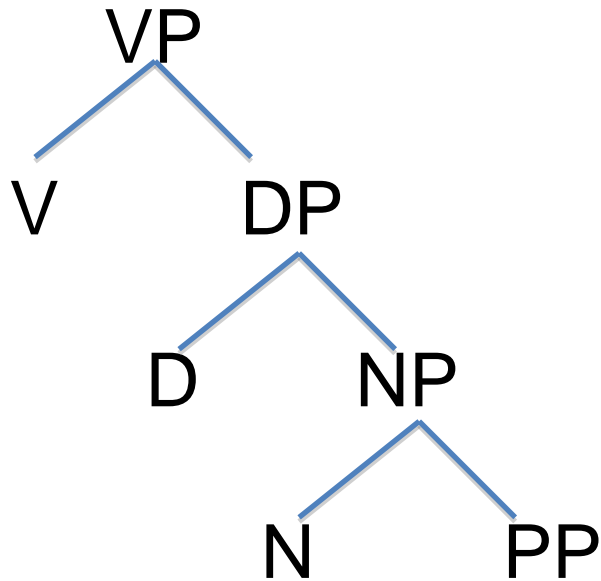
## ADDITIONAL PROBLEMS, I

A similar problem arises when the nominal idiom is selected by a verb which is also part of the idiom.



Here, no matter which analysis of the status of the determiner we opt for, both hypotheses on idioms formation do converge in their predictions.

## ADDITIONAL PROBLEMS, II



*Theory I. (Constituency):* the only constituent including V, D and N also contains PP: the complement needs to be included in the idiom.

*Theory II. (Selection):* V selects N (or selects D that selects N), so the argument of N must be included in the idiom.

## ADDITIONAL PROBLEMS, III

However the prediction proves to be wrong at the light of cases like the following:

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

(7) Etre à deux pas de X  
*Lit.* To-be at two steps from  
'To be very close to...'

(8) Essere a un passo da X  
*Lit.* To-be at a step from  
'To be very close to..'

## OUR LINE OF ATTACK: LATE MERGE

The hypothesis that we will propose here is that nouns do not take complements, and this is why idioms of this sort exist. Nominal modifiers are not selected, they are adjuncts, hence they can be late merged.

For example in (4), we assume that the PP can be inserted after D and N have merged.

(4) *du fili (di spaghetti, di carbonara, di matriciana, al pesto, di insalata, di X (food with string-like shape)*

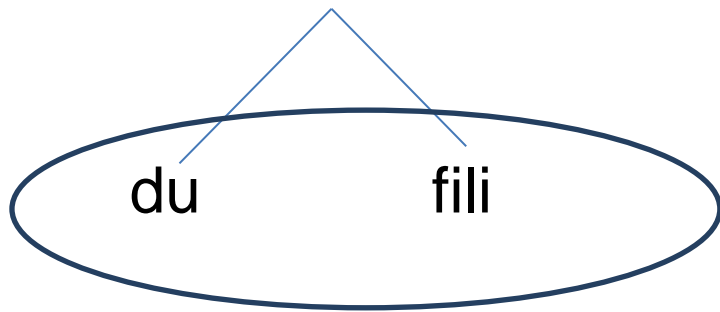
*Lit.* Two strings of...

‘A portion of X’

# LATE MERGE

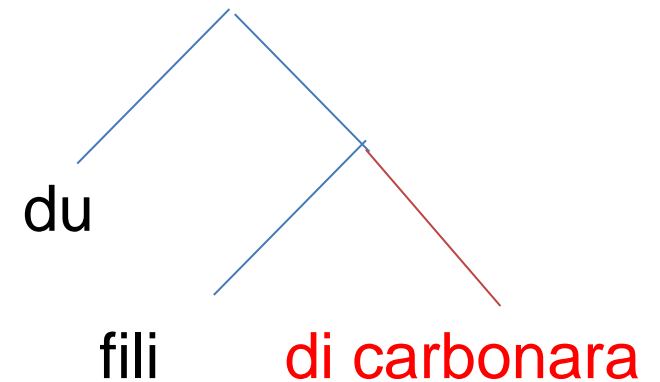
'Di carbonara' is late merged after merge of 'du'. The circled area indicates the constituent that gets the idiomatic reading.

## INITIAL STRUCTURE



{du, fili}

## STRUCTURE AFTER LATE MERGE



{du, {fili {di carbonara}}}

## WHEN THE PP FOLLOWING THE NOUN DOES ENTER THE IDIOM

Of course the PP modifier of the noun *can* be part of the idiom.

(9) Lui è una testa di cazzo

He is a head of dick

'He is a dickhead'

(10) Ha fatto un buco nell'acqua

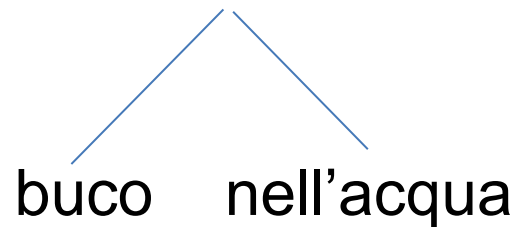
(he) made a hole in the water

'He failed/he went nowhere'

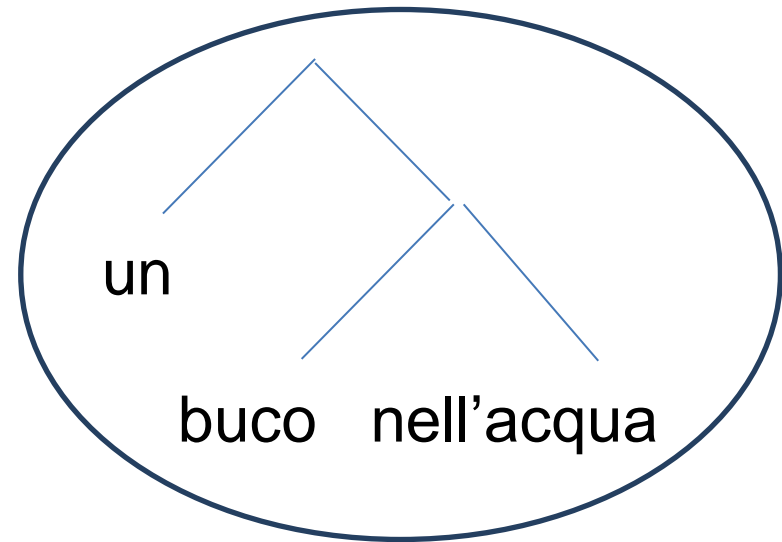
We explain this because late merge is *possible*, not forced.

# EARLY MERGE

'Nell'acqua' is early merged (before merge of 'un'). The circled area indicates the constituent that gets the idiomatic reading.



{buco, nell'acqua }



{un, {buco {nell'acqua}}}



# OUR APPROACH AND THE THEORY OF IDIOMS FORMATION

The approach in terms of late merge is a revised form of the constituency theory, with the important twist that constituency is dynamically defined. In particular, constituents that precede late merge of adjuncts can undergo idiom formation.

# THREE QUESTIONS FOR THE DYNAMIC VERSION OF THE CONSTITUENCY THEORY OF IDIOMS

Three questions immediately arise:

1. Is the assumption that the PP that follows the noun is an adjunct independently motivated?
2. Can the late merge approach explain the facts that go against the traditional version of the constituency theory?
3. What factors determine early/late merge?

# **1. SUPPORTING THE HYPOTHESIS THAT NOUNS DO NOT TAKE COMPLEMENTS**

# NOUNS DO NOT TAKE COMPLEMENTS

We have devoted a large part of Cecchetto & Donati (2015) to argue for the hypothesis that nouns do not take the complements the way verbs do and therefore so-called PP complements of nouns can be late inserted (see also Adger 2014 for independent evidence).

Cecchetto, C. & Donati C. (2015). *(Re)labeling*. Cambridge Mass: MIT Press.

Since most of the arguments in Cecchetto & Donati (2015) are not relevant for the present discussion about idioms we briefly summarise only two very general considerations supporting this conclusion.

## THETA CRITERION EXEMPTION

Even so-called complements of nouns are never required for the structure to be acceptable, unlike the complements of transitive verbs. This is usually expressed by exempting the nouns from the theta criterion, but this is a tacit way to 'adjunctivize' the so-called complement of the noun.

## CONSTITUENCY TESTS, CLAUSAL DOMAIN

In the clausal domain, proform substitution indicates that the verb + direct object form a minimal constituent excluding the subject:

(11) John bought a house and Mary did that too.

However, a proform that replaces subject + verb excluding the direct object is cross-linguistically unattested. In no language is a sentence like (12) grammatical.

(12) \*John bought a house but did that no car  
(Intended meaning: John bought a house but not a car)

Observations of this kind motivate Baker's (2009) Verb-Object Constraint, one of the best candidates for a language universal.

## CONSTITUENCY TESTS, NOMINAL DOMAIN

However, a proform can replace the unit formed by D+N, crucially excluding the alleged complement of the noun.

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

This is consistent with the hypothesis that there is a moment in the derivation in which D+N is a constituent that excludes the PP, because PP has not been merged yet.

We assume that the same dynamic theory of constituency formation that explains idiom formation can explain pro-form substitution. A proform can stand for the temporary constituent that includes D + N to the exclusion of the PP.

## **2. VARIABILITY OF THE DETERMINER**



## WHEN THE DETERMINER CHANGES INSIDE THE IDIOM

The variability of the determiner in idioms like (1) is still somehow challenging for a theory based on constituency.

(1)

Pull *some* strings

pull *a few* strings

pull *yet more* strings

One way to go: since the verb selects the noun (not the determiner) also the determiner can be late merged. This recovers the fundamental part of Bruening's intuition by rephrasing it in terms of late merge (elements that are not selected are adjunct-like and can be late merged).

### **3. WHAT FACTORS DETERMINE TIMING OF MERGE?**

# CONSTRAINTS ON TIMING OF MERGE

1. It is not semantic: even constituents that would be analyzed as complements from a semantic point of view can be late merged (e.g. *trente siz façons de X*).
2. It is constrained by structural conditions: in Cecchetto and Donati (2015) we show that the same modifier that is late merged in a relative clause is early merged in an interrogative clause, and we argue that this explains some well-known Reconstruction riddles.
  - (14) a. The professor of John's that he always praises
  - b. \*Which professor of John's does he always praise? (Munn 1994; Safir 1999)

# OPEN ISSUES

## ONLY NUMERALS?

The cases of unexpected idioms discussed up to now all involve numerals (“due fili”, “quattro passi”, “trente-six façons” etc.). It can be claimed that numerals cannot be assimilated to other types of determiners. This would *not* impact on our argument because, even if numerals are not real Ds, they still are not expected to enter an idiom together with N with the exclusion of the PP that follows N.

However, the question whether the unexpected pattern of idiom formation (from now “weird idioms”) is restricted to numerals is legitimate.

## WEIRD IDIOMS WITHOUT NUMERALS?

In the following idioms, arguably the D is the indefinite article (although in Italian the indefinite article is homophonous with the numeral ‘one’) or a demonstrative.

(15)

quella cazzo di (penna, prof, X)  
uno straccio di (prova, lavoro, X)  
un casino (di mail, gente, posta, , X)  
una pioggia di X  
una palla di X  
un mare di X  
un pezzo di X  
un mondo di X  
una montagna di X

*lit.* that cock of (‘the fucking X’)  
*lit.* a rug of X (‘no X’)  
*lit.* a mess of X (‘a lot of X’)  
*lit.* a rain of (‘lots of X’)  
*lit.* a ball of (‘a boring X’)  
*lit.* a sea of (‘a lot of X’)  
*lit.* a piece of X (an outstanding X)  
*lit.* a world of X (a lot of X)  
*lit.* a mountain of X (a mountain of X)

In all these idioms, D+N have a fixed idiomatic meaning but the PP is variable.

## AN ILLUSTRATIVE EXAMPLE: “UNO STRACCIO DI”, I

Still, these cases are less clear than those with numerals, since the determiner introducing the idiom is not totally fixed. We show this with the NPI ‘uno straccio di X’ (*lit.* a rag of X).

- (16) a. Non c’è uno straccio di prova contro di lui  
*Lit.* Not there is a rag of evidence against of him  
There is no evidence whatsoever against him
- b. Non trova uno straccio di marito  
*Lit.* (she) not finds a rag of husband  
She cannot find anyone to marry
- c. Un mare senza uno straccio di sabbia  
*Lit.* A sea without a rag of sand  
A seaside with no beach

## AN ILLUSTRATIVE EXAMPLE: “UNO STRACCIO DI”, II

In negative contexts the determiner introducing ‘straccio’ (rag) is fixed, namely it must be the indefinite article:

- (17) a. \*Non c’è quello/lo straccio di prova contro di lui  
*Lit.* Not there is that/the rag of evidence at his
- b. \*Non trova quello/lo straccio di marito  
*Lit.* (she) not finds that/the rag of husband
- c. \*Un mare senza quello/lo straccio di sabbia  
*Lit.* A sea without that/the rag of sand



## AN ILLUSTRATIVE EXAMPLE: “UNO STRACCIO DI”, III

However, in another syntactic context, the determiner introducing ‘straccio’ (rag) may change:

(18) Non ce la fa più con quello straccio di marito che si ritrova

*Lit.* (S)he not there does any-more with that rag of husband that (s)he herself/himself finds!

(S)he cannot handle the poor husband (s)he has.

Notice that (17) crucially involves a relative clause.

## ANOTHER ILLUSTRATIVE EXAMPLE: “UN MARE DI”

The D introducing the idiomatic reading of ‘mare’ (sea) is the indefinite article (19a-b). It can change only if a relative clause is involved (19c):

- (19) a. Ho mangiato un mare di biscotti  
*Lit.* (I) have eaten a sea of biscuits  
I ate a lot of biscuits
- b. \*Ho mangiato quel/il mare di biscotti  
*Lit.* (I) have eaten that/the sea of biscuits
- c. Ha mangiato quel mare di biscotti che Maria aveva preparato  
*Lit.* (he) has eaten that sea of biscuits that Maria had prepared

## DETERMINERS' VARIABILITY

Given this situation one cannot totally exclude that the idiomatic reading is associated just to the noun *straccio* or *mare*. However, this would predict a bigger variability in the determiner choice than it is actually observed.

We suspect that the relative clauses allow more determiner variability, because they actually involve two determiners for the head noun, one internal to the relative and one external (see Bianchi 1999, Bhatt 2002, Kayne 1994, Cecchetto and Donati 2015).

## CONCLUSIONS, I

We have identified a weird idiom configuration in which D+N enters the idiomatic reading while the PP that follows the N does not. This is unexpected under current syntactic theories of idiom formation but can be explained by:

- (i) assuming a late merge approach to adjuncts
- (ii) assuming that even so-called complements of the noun are adjunct-like categories
- (iii) assuming a *dynamic* constituency theory to idiom formation.

## CONCLUSIONS, II

If we are on right track, there is a fundamental difference between the verbal and the nominal domain. In the verbal domain, idioms of the form Subject + Verb to the exclusion of the direct object are impossible because verbs take real complements, while in the nominal domain idioms of the form D + N to the exclusion of the PP selected by the noun are possible because nouns do not take complements.

This is one of the many pieces of evidence that call for a revision of the traditional hypothesis of full parallelism between clausal and nominal structure.

## REFERENCES

- Bhatt, R. 2002. The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification. *Natural Language Semantics* 10: 43–90.
- Bianchi, V. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.
- Bruening, B. 2010. Ditransitive asymmetries and a theory of idiom formation, *Linguistic Inquiry* 41: 519-562.
- Cecchetto, C. & C. Donati. 2010. On labeling: Principle C and head movement. *Syntax* 13: 241-278.
- Cecchetto, C. and Donati C. 2015. *(Re)labeling*. Cambridge Mass: MIT Press.
- Donati, C. & C. Cecchetto. 2011. Relabeling Heads: A Unified Account for Relativization Structures. *LI* 42: 519-560.
- Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Munn, A. 1994. A minimalist account of *reconstruction* asymmetries. In *Proceedings of NELS 24*. GLSA.
- O'Grady, W. 1998; The syntax of Idioms. *Natural Language and Linguistic Theory* 16: 279-312.
- Safir, Ken. 1999. Vehicle change and *reconstruction* in A-chains. *Linguistic Inquiry* 30:5 87-621.