Empirical Approaches to Elliptical Constructions

Class 3: Gapping in Romance

Gabriela Bîlbîie\textsuperscript{1} and Anne Abeillé\textsuperscript{2}
\textsuperscript{1}University of Bucharest
gabriela.bilbiie@gmail.com
\textsuperscript{2}Université Paris Diderot-Paris 7
anne.abeille@linguist.univ-paris-diderot.fr
LSA 2017 Linguistic Institute

13 July 2017, University of Kentucky
Parallelism constraints
Evidence for a fragment-based analysis
Gapping vs. Stripping

Class 3 – Content

1 Parallelism constraints

2 Evidence for a fragment-based analysis

3 Gapping vs. Stripping
Disproportionate studies on gapping


- Gapping studies on other languages: Dutch (Neijt 1979, Aelbrecht 2009), Chinese (Paul 1999, Tang 2001, Ruixi Ressi 2008), Russian (Kazenin 2001, Agafonova 2014), Turkish (Ince 2009), Farsi (Farudi 2013), Ancient Greek (Gaeta & Luraghi 2001), Latin (Panhuis 1979); 'exotic' languages such as Quechua (Pulte 1971), Zapotec (Rosenbaum 1977).

- However, gapping has by and large been overlooked in Romance. Here, we mostly concentrate on two Romance languages: French and Romanian. Our generalizations are supported by attested data from:
  - ESTER Copus (Gravier et al. 2004): radio broadcast news.
Advantages of studying gapping in Romance

Romance languages allow us to better observe and understand the behaviour of gapping construction.

- **Parallelism constraints** (so much discussed in the literature!): more empirical evidence against strict syntactic parallelism between the target clause and the source.

- **Competitive syntactic analyses**: more empirical evidence in favour of a construction-based analysis of gapping (and against deletion or movement-based analyses).
Plan

1. Parallelism constraints
2. Evidence for a fragment-based analysis
3. Gapping vs. Stripping
Strong syntactic parallelism?

- Standard assumption: Strong syntactic parallelism should hold between the target clause and its source (Hartmann 2000).
  - 'Connectivity effects' (case marking, preposition marking, category identity, same number of remnants and correlates, same word order): strong argument for structural approaches.

- Assumption first challenged by Sag et al. (1985):
  - The order of remnants in the gapped clause does not necessarily need to parallel that of their correlates in the source clause, provided that this order is licit in the grammar (1-a).
  - Remnants may differ from their correlates with respect to their category (1-b).
  - Still, syntax is not left completely unconstrained: each remnant is required to match some subcategorization frame of the verbal predicate its correlate depends on.

(1) a. A policeman **walked** in at 11, and at 12, a fireman. (Sag et al. 1985)
   b. Pat has become [crazy]_{AP} and Chris [an incredible bore]_{NP}.
   c. *Pat has become [crazy]_{AP} but Chris [in good spirit]_{PP}.
   d. He became \{crazy | an incredible bore | *in good spirit\}. 
In other words, remnants and their correlates in gapping constructions obey the same syntactic constraint than conjuncts in ordinary constituent coordinations: each must match some subcategorization of the shared predicative material, though not necessarily the same one.

(2) a. He has become [crazy]_{AP} and [an incredible bore]_{NP}.
   b. *He has become [crazy]_{NP} but [in good spirit]_{PP}.

French and Romanian parallel English in this respect: order and category asymmetries are allowed, provided the syntactic constraint mentioned above is observed.
Category asymmetries

French examples:

(3)  a. Certains réclament [des augmentations]_{NP}, d’autres [qu’on leur garantisse la sécurité]_{S}.
    'Some demand some increases, others (demand) that one warrants their safety.'
  b. *Certains réclament [des augmentations]_{NP}, d’autres [être mieux protégés]_{VPinf}.
    'Some demand some increases, others (demand) to be better protected.'
  c. Ils réclament {des augmentations | qu’on leur garantisse la sécurité | *être mieux protégés}.
    'They demand {some increases | that one warrants their safety | to be better protected}.'

(4)  a. Ils réclament [des augmentations]_{NP} et [qu’on leur garantisse la sécurité]_{S}.
    'They demand some increases and that one warrants their safety.'
  b. *Ils réclament [des augmentations]_{NP} et [être mieux protégés]_{VPinf}.
    'They demand some increases and to be better protected.'
In Romanian, remnants may differ from their correlates not only with respect to their basic category (5-a), but also (in some cases) with respect to case marking; in (5-b) the second remnant *tuturor copiilor* (‘to all the children’) bears an affix marking a dative form, while its correlate in the source *la trei dintre copii* (‘to three of the children’) is introduced by a prepositional marker *la* (‘to’) asking for an accusative form.

\[(5) \quad \text{a. } \underline{\text{Ion citeşte [ziua]}}_{NP}, \text{ iar Maria [pe-ntuneric]}_{PP}. \]
\[\text{‘Ion reads during the day and Maria at night.’}\]

\[\text{b. } \underline{\text{Ion oferă mere \[la trei dintre copii\]}_{PP}, \text{ iar Maria [tutu} \underline{\text{ror copiilor]}_{NP}.}} \]
\[\text{‘Ion gives apples to three among children and Maria all.} \underline{\text{DAT children.} \text{DEF.DAT}}\]
\[\text{‘Ion gives apples to three of the children, and Maria to all of them.’}\]
Word order asymmetries

Romance languages provide more striking asymmetries, as mainly shown by Romanian data below. Given its relatively free word order, there are many ways to linearize remnants and correlates in Romanian.

(6) a. Dimineața (EU) spăl (EU) vesela (EU), iar in.the.morning (I) wash (I) the.dishes (I), and seara IOAna. in.the.evening Ioana ‘In the morning I wash the dishes and in the evening Ioana does.’

b. EU spăl vesela dimineața, iar seara IOAna. I wash the.dishes in.the.morning and in.the.evening Ioana

c. Eu spăl vesela dimiNEAța, iar Ioana SEAra. I wash the.dishes in.the.morning and Ioana in.the.evening

d. DimiNEAța spăl eu vesela, iar Ioana SEAra. in.the.morning wash I the.dishes and Ioana in.the.evening
Asymmetries and subject pro-drop

- The number of remnants may differ from the number of realized correlates, as shown by the subject pro-drop phenomenon for Romanian (7-a) and Italian (7-b).

- Lack of empirical data on gapping with pro-drop. Some people consider such examples as unacceptable.

(7)  

a. Lunea **merg** la film, iar **sora mea** la muzeu.
   'On Monday, I go to the cinema, and my sister goes to the museum.'

b. Mangio la **pasta e Giovanni il riso**.
   'I eat pasta and Giovanni eats rice.'
We set up an experiment on acceptability judgments in order to check if prodrop is really natural and acceptable with gapping in Romanian.

Participants: 29 Romanian native speakers from age 19 to 21, 1st and 2nd year students at the University of Bucharest (Faculty of Foreign Languages).

A crossed factorial design (2x2 design) with two factors (structure and subject type), each with two levels. Each level of factor 1 is combined with each level of factor 2. ⇒ 4 conditions

- Values for structure: gapped clause vs. full clause
- Values for subject type: overt subject vs. prodrop in the source clause

A within-subjects design: each participant sees every condition.
Experimental items (4 conditions for each item):

- Condition 1: **gapping** and **overt subject**
- Condition 2: **gapping** and **prodrop**
- Condition 3: **full clause** and **overt subject**
- Condition 4: **full clause** and **prodrop**

(8) Ne-am făcut deja planul pentru weekend.
'We have already planned our weekend.'

a. **Eu** voi merge la film, iar **sora mea** la muzeu.
I will.1sg go to cinema and sister.DEF POSS to museum
'I will go to the cinema, and my sister to the museum.'

b. **Voi** merge la film, iar **sora** **mea** la muzeu.
will.1sg go to cinema and sister.DEF POSS to museum

b. **Eu** voi merge la film, iar **sora mea** va merge la muzeu.
'I will go to the cinema, and my sister will go to the museum.'

d. **Voi** merge la film, iar **sora** **mea** **va** merge la muzeu.
will.1sg go to cinema and sister.DEF POSS will go to museum.
- Only dynamic and agentive verbs (e.g. to enter, to go out, to arrive, to leave, to come, to run, to flight, to dance, to work, to manifest, to ski, to work...).

- Only future tense is used in the experiment (in order to avoid underspecified verbal forms).

- The postverbal constituent is always a locative or temporal complement or adjunct (e.g. to the cinema vs. to the museum, at 3p.m. vs. at 5p.m.).

- Only different persons (e.g. no 3rd person in both clauses).

- Control items: truly unacceptable sentences, with wrong verb form, wrong agreement, etc.

(9) *Fetele doarme sus, iar băieţii dormi jos.  
     girls.DEF sleep.3SG upstairs and boys.DEF sleep.2SG downstairs  
     'The girls are sleeping upstairs and the boys are sleeping downstairs.'
- The material contains 24 experimental items and 12 control items.
- Each participant has to give 6 judgments (observations) per condition.
- 4 groups of participants, 4 questionnaires.
- Each questionnaire includes 24 experimental items randomized with 12 control items and 19 fillers (from another experiment, in order to hide the experiment’s purpose and to balance types of items). ⇒ 55 observations per questionnaire.
- Likert scale: Participants judge each sentence individually along a numerical scale (5-point scale, with a middle point, in order to give participants the option of saying that they don’t know whether the sentence should fall on the acceptable or unacceptable side of the spectrum).
  - 1 = not natural at all vs. 5 = completely natural.
Results:

- The average judgments in each condition (≈ mean rates):
  - condition 1 (+gapping, -prodrop): 4.694444444
  - condition 2 (+gapping, +prodrop): 4.611111111
  - condition 3 (-gapping, -prodrop): 4.638888889
  - condition 4 (-gapping, +prodrop): 4.628571429
  - control items: 1

- There is no significant difference between gapping with and without prodrop; prodrop is indeed acceptable with gapping.

- There is no significant difference between the use of gapping and the use of the full clause. Need of further experimentation!

- These results support the assumption that the syntactic parallelism in gapping is not as strong as assumed.
Asymmetries and object pro-drop

- Same asymmetries can apply with object pro-drop, as shown for Romanian (remnants may correspond to a weak element, such as an adverbial (10-a) or pronominal (10-b) clitic) and French (the direct object remnant has no explicit correlate in the source, as in (11)).

(10) a. Dan tot mai citește, dar prietena lui absolutely
nothing
'Dan still reads, but his girlfriend (doesn't read) anything at all.'

b. Ion mi-e prieten, iar tîie dușman.
'Ion is my friend and your enemi.' ('For me, Ion is a friend, and for you, he is a enemi.')

(11) a. Marie nage bien, mais Paul seulement la brasse.
'Marie swims well, but Paul only the strokes.'

b. Paul boit trop, mais son frère que de l’eau.
'Paul drinks a lot, but his brother only water.'
In some cases, remnants have not even an implicit available correlate in the source.

(12) a. [Ion] a cumpărat [un ziar], iar [Maria] [o jucărie] pentru fetița ei.
    ’Ion has bought a newspaper, and Maria a toy for her daughter.’

b. [Dan] merge [la munte], iar [Maria] probabil [la mare].
    ’Dan goes to the mountains, and Maria probably at the seaside.’

We conclude that syntactic parallelism operates neither at the level of phrase structure, nor at the level of word order, but rather at the more abstract level of grammatical functions, as listed in the argument structure of predicates.
Discourse parallelism

While syntactic parallelism is not strict, discourse parallelism is clearly required. As it is the case for English (cf. Kehler 2002), some symmetric relation (viz. preserved when the order of the conjuncts is reversed) must hold between the source clause and the gapped clause in French and Romanian.

- Gapping is felicitous with relations such as parallelism or contrast.
- It is excluded with cause-effect relations, such as concession or condition.

This immediately explains why gapping occurs neither in comparative correlatives (13-a) (which involve an if...then... interpretation, cf. Beck 1997), nor in causal coordinations (13-b) or subordinated contexts (13-c).
(13) a. Plus Marie lira de romans et plus Jean
    the.more Maria read.FUT INDEF novels and the.more Jean
    *(lira) de BD.
    read.FUT INDEF comics
    'The more novels Marie will read, the more comics Jean will (read).'

b. Jean a mis un costume, car Marie *(a mis) une jolie
    Jean has put a suit because Marie has put a nice
    dress
    'Jean dressed in a suit, because Marie dressed in a nice dress.'

c. Jean voit une autre fille, parce que sa copine *(voit) son meilleur
    Jean dates an other girl because his girlfriend dates his best friend.'
Semantic parallelism

- For symmetry to hold, each remnant must stand in semantic contrast with respect to a correlate in the source, as discussed by Sag (1976), and more recently by Hartmann (2000) and Repp (2009). An appropriate contrast can only be established between elements of a well-defined alternative set (different agents, different locations, different times, etc.).

- Evidence from Romanian:
  
  - Romanian distinguishes itself from French and other Romance languages in this respect by having a special ‘contrastive’ conjunction *iar* (‘and’), which is the most used conjunction in gapping coordinations, as in Slavic languages which have equivalent connectives, e.g. *a* in Russian and Polish (Jasinskaja & Zeevat 2009, Repp 2009).

(14) a. Ion **adoră** fotbalul, **iar** Maria baschetul. (Romanian)
   ‘Ion likes football, and Maria likes basketball.’

b. Oleg **ljubit** futbol, **a** Maria basketbol. (Russian)
   ‘Oleg likes football, and Maria basketball.’
Why this affinity of \textit{iar} for gapping in Romanian?

- This conjunction meets all the constraints imposed by gapping constructions.
  - Double contrastiveness constraint; \textit{iar} requires the presence of two contrastive pairs, i.e. two pairs of elements such that an element of each pair is present in each conjunct, and such that the elements in each pair are contrastive: they must belong to the same domain and be different (Zeevat 2004). Gapping typically has two contrastive pairs too.
  - The contrastive pairs linked by \textit{iar} provide the answers to an implicit multiple question. Gapping is much more natural in answer to multiple \textit{wh}-questions (Steedman 2000).
  - Gapping is reserved to symmetric discourse relations (in particular, parallelism and contrast, cf. Levin & Prince 1986, Kehler 2002); the conjunction \textit{iar} is compatible only with this kind of relations.
  - Gapping applies only at the clausal level; \textit{iar} links only clauses or fragments with a propositional content.
Plan

1. Parallelism constraints

2. Evidence for a fragment-based analysis

3. Gapping vs. Stripping
Three kinds of analyses have been explored to account for this unusual mapping form/meaning:

1. **Movement-based analysis**, rejecting ellipsis, appeals to some leftward movement process that combines across-the-board extraction of the shared head verb out of each conjunct and asymmetric extraction of non-shared constituents preceding the head verb out of the first conjunct (Johnson 1994, 1996, 2009).


3. **Construction-based analysis**, appealing to a dedicated meaning-form rule, namely a ‘construction’, that maps a headless structure to a clausal meaning (see a.o. Sag et al. 1985; Chao 1987; Steedman 1990; Gardent 1991; Culicover & Jackendoff 2005).

(15) Paul viendra lundi et Jean mardi.

Paul come.FUT Monday and Jean Tuesday
‘Paul will come on Monday, and Jean on Tuesday.’
According to the movement-based analysis, gapping does not result from ellipsis, but rather from some movement process that extracts the head verb 'across-the-board', namely out of each conjunct. Subclausal coordination.

Some additional movement rule must be posited that extracts several non-shared constituents belonging to the first conjunct (i.e. the subject phrase), which linearly precede the main verb.
Semantic motivation of the movement


(16) a. No woman$_i$ can join the army and her$_i$ girlfriend the navy.
   b. *No woman$_i$ can join the army and her$_i$ girlfriend can join the navy.

(17) Mrs. Smith can’t dance or Mr. Smith sing.
   a. = Mrs. Smith can’t dance and Mr. Smith can’t sing.
   b. ̸= Mrs. Smith can’t dance or Mr. Smith can’t sing.

(18) Not every girl ate a GREEN banana and her mother a RIPE one.
Parallelism constraints  Evidence for a fragment-based analysis  Gapping vs. Stripping

Problems for movement-based analysis

This analysis does not immediately extend to naturally occurring examples, where the gap includes not only the head verb, but also various elements which do not form continuous strings of words (see PTB data from class 2, and Huddleston et al. 2002 for similar examples in English).

(19)  

a. Pour un salaire de 20 000F, le coût d’emploi serait majoré de 300F par mois et [pour un salaire de 50 000F, increased by 300F per month and for a wage of 50 000F de 1 500F]. (FTB) by 1 500F

’For a wage of 20 000F, the employment cost would be increased by 300F per month, and for a wage of 50 000F, it would be increased by 1 500F.’

b. Jean pense que la France va gagner et [Marie Jean thinks that DEF France will win and Marie l’Argentine].

DEF-Argentina

‘Jean thinks that France will win, and Marie thinks that Argentina will win.’
Moreover, it wrongly predicts that initial conjunctions, which arguably mark the left edge of the first conjunct in Romance (cf. Mouret 2005, 2007; Bîlbîie 2008), should be realized after the alleged moved material, and not before.

(20)  
\begin{enumerate}
\item \textbf{Ou bien} Paul \underline{dormira} chez Marie \textbf{ou bien} Marie chez Paul.
\begin{itemize}
\item \textit{either Paul will.sleep at Marie or Marie at Paul}
\end{itemize}
\textit{'Either Paul will sleep by Marie’s or Marie by Paul’s.'}
\item *Paul \underline{dormira} \textbf{ou bien} chez Marie \textbf{ou bien} Marie chez Paul.
\begin{itemize}
\item Paul will.sleep \textit{either at Marie or Marie at Paul}
\end{itemize}
\end{enumerate}

(21)  
\begin{enumerate}
\item \textbf{Ni} le compromis \underline{ne} \underline{me} \underline{paraît} justifié, \textbf{ni}
\begin{itemize}
\item neither DEF compromise NEG me seems justified, nor
\end{itemize}
\begin{itemize}
\item l’acceptation pure et simple nécessaire.
\end{itemize}
\textit{DEF-acceptance pure and simple necessary}
\textit{‘Neither the compromise seems to me justified, nor the pure and simple acceptance seems to me necessary.’}
\item *Le compromis \underline{ne} \underline{me} \underline{paraît} \textbf{ni} justifié, \textbf{ni}
\begin{itemize}
\item DEF compromise NEG me seems neither justified, nor
\end{itemize}
\begin{itemize}
\item l’acceptation pure et simple nécessaire.
\end{itemize}
\textit{DEF-acceptance pure and simple necessary}
The movement-based analysis (understood as a sub-clausal coordination) doesn't account for the huge frequency of the clausal coordinator *iar* in Romanian gapping; this conjunction only applies to the clausal level, being excluded with sub-clausal constituents (Bîlbîie & Winterstein 2011).

⇒ We conclude that movement creates more problems than it solves.
Deletion-based analysis

- Deletion-based analyses of gapping and other related verbal ellipsis come in two varieties:
  - Classical accounts (Sag 1976) assign to the gapped clause the syntactic structure of an ordinary sentence with some material including the head verb left unpronounced.
  - More recent accounts within derivational frameworks furthermore consider remnants to be extracted in some functional projections in the left periphery (Coppock 2001).
Problems for extraction-based accounts

Violation of locality constraints. As Culicover and Jackendoff (2005) observe, remnants do not obey island constraints in English (i.e. they can be embedded under a propositional domain), contrary to what has been claimed since Ross (1967). When more than the head is deleted, remnants can appear in what would be an island for extraction, for example a circumstantial adjunct or a relative clause.

(22) a. Robin knows a lot of reasons why dogs are good pets, and Leslie cats. (Culicover & Jackendoff 2005)
   b. Robin believes that everyone pays attention to you when you speak French, and Leslie, German.

(23) a. *[Which animals]; does Robin know a lot of reasons why –i are good pets?
   b. *[Which language]; does Robin believe that everyone pays attention to you when you speak –i?
(24)  

a. In the past, it has been the husband who has been dominant and the wife passive. (brwn-21990)

b. Bo decided who is working tomorrow, and Mia, the next day. (Chaves 2005)

c. That’s not how I remember it at all. And yet he cannot be the one who’s correct, and everyone else – millennia of people – wrong. (Hanya Yanagihara, A Little Life, p. 321-322)
Similar data obtain in Romance (cf. French data below): remnants may occur not only in circumstantial adjuncts (25-a) and relative clauses (25-b), but also in infinitival subjects (25-c), out of which nothing can be extracted.

(25)  

a. **Quand tu parles** chinois, tout le monde **t’admire**, mais anglais personne.  
   ‘When you speak Chinese, everyone admires you, but (when you speak) English, nobody (does).’
   
b. **C’est Paul qui fait** la vaisselle et Marie la lessive.  
   ‘It’s Paul who does the dishes and Marie the washing.’
   
c. **Comprendre le texte traduit** c’est laborieux et le texte original encore plus laborieux.  
   ‘Understanding the translated text is painstaking and (understanding) the original text is all the more painstaking.’

⇒ Extraction of remnants in the left periphery is therefore not empirically supported. As a result, deletion, if it is adopted as the source of gapping, must be allowed to target non-constituent strings.
Problems for deletion accounts without extraction

- Syntactic mismatches (different agreement, different clitics, absence of correlates).

- Pronominal clitics hosted by the source verb may differ from those required when the alleged deleted verb is reconstructed.

(26) a. Paul en a lu seulement certains, mais Max (\{les | *en\}
Paul CL has read only some but Max \{CL.3PL | CL\}
a lu) presque tous.
has read almost all
‘Paul has read only some of them, but Max has read almost all of them.’

b. Paul les a lus, vos livres, et Max (\{en | *les\}
Paul CL.3PL has read.M.PL your books and Max \{CL | CL.3PL\}
a lu) seulement certains.
has read only some
'Paul has read your books, but Max has read only some of them.’
Object Clitic Doubling in Romanian:

(27) a. Ion l-a văzut pe Dan, iar Ana (a Ion cl.masc-has seen dom Dan and Ana (has văzut-o) pe Maria.
seen-cl.fem) dom Maria
Ion has seen Dan, and Ana Maria.’

b. Maria le-a citit pe toate, dar Ana
’Maria cl.fem.3pl-has read dom all but Ana
(*le-) (a citit) doar câteva.
cl.fem.3pl (has read) only some
‘Maria has read them all, but Ana only some.’
Moreover, while absent in the source, some clitics, such as negative particles or subject weak pronouns, may be required in the reconstructed clause given the form of remnants.

- The n-word *aucun* (‘none (of them)’) requires for example the particle *ne* (‘not’) on the reconstructed verb in French (28-a).
- The strong pronoun *moi* (‘me’) is not easily used as subject in the reconstructed form, unless it is doubled by the pronominal clitic *je* (‘I’) as in (28-b).

(28) a.  
  
  Paul en a lu peu, et Marie (*\(n\')en a lu) absolument  
  Paul CL has read few and Marie NEG-CL has read absolutely  
  aucun.  
  none  
  ‘Paul has read few of them and Marie has read absolutely none of  
  them.’

b.  
  
  Jean aime les pommes et moi (*\(j\')aime) les oranges.  
  Jean likes DEF apples and me I-like DEF oranges  
  ’Jean likes apples and I like oranges.’

⇒ Identity conditions required as a trigger for deletion cannot be stated in a simple way.
Gapped clauses do not have the same distribution as their complete counterpart. As noted by Culicover and Jackendoff (2005), there are contexts where gapped clauses are allowed and complete clauses excluded.

- A gapped clause may be introduced by functors such as constituent negation adverbs or lexicalized comparative connectives, which do not combine with finite sentences.

b. Robin speaks French as well as Leslie (*speaks) German.

(30) a. Bill invited Jane and not Jane (*invited) Bill.
b. Bill wanted to meet Jane as well as Jane (*wanted to invite) him.

(31) a. Paul dormira chez Marie et non pas Marie (*dormira) chez Paul.
   ‘Paul will sleep by Marie’s and not Marie by Paul’s.’
b. Paul a cueilli des framboises ainsi que Marie (*a cueilli) des fraises.
   ‘Paul picked raspberries as well as Marie strawberries.’

⇒ Syntactic reconstruction is not always warranted. Such data stand as a serious challenge for any deletion-based account.
A fragment-based analysis

- These empirical data immediately fall out from a construction-based analysis (with semantic reconstruction of ellipsis) which does not derive the unusual meaning/form mapping in the gapped clause from hidden syntactic structure.

- If there is no head verb in the gapped clause, then the gapped clause itself is not finite and thus appropriate for combination with functors selecting some non finite constituent.
A non-structural approach (cf. Ginzburg & Sag 2000): a 'what you see is what you get' syntactic structure for gapping constructions.

- Cf. Simpler Syntax Hypothesis: "The most explanatory syntactic theory is one that imputes the minimum structure necessary to mediate between phonology and meaning." (Culicover & Jackendoff 2005)

- Semantic reconstruction: Fragments' content is built from the meaning of the source, the remnants and their correlates by some relation $R_{sem}$.

- Discursive constraint: Some symmetric discourse relation must hold between conjuncts.

- The only syntactic constraint: each remnant must match a possible subcategorization of the verbal predicate its correlate depends on. Otherwise, they may differ from their correlates with respect to their category, their position or their surface realization.

- The overall construction is a particular type of asymmetric coordination with the main conjunct as being non elliptical and verbal, and the gapped one as fragmentary and non verbal.
All the core properties of gapping can be easily represented within a construction-based framework that relies on rich inheritance hierarchies of typed feature structures, such as Head-driven Phrase Structure Grammar (Ginzburg & Sag 2000).

Nothing hinges on that particular choice: our analysis could as well be incorporated within the recent Sign-Based Construction Grammar framework advocated by Ivan Sag and colleagues (see Boas & Sag 2012, Sag 2012) or within the Berkeley Construction Grammar framework of Fillmore and Kay (1996).
Plan

1. Parallelism constraints
2. Evidence for a fragment-based analysis
3. Gapping vs. Stripping
The same construction or not?

- Is Stripping (also called Bare Argument Ellipsis, see class 1) a subtype of Gapping or not?
  - Difference in the number of remnants: usually, one in Stripping vs. two in Gapping constructions.
  - The single remnant is very often combined with an adverb such as *too* (32-a) or *not* (32-b).

(32)  
- a. John \underline{drinks} scotch, \[and Bill \underline{too}\].
- b. John \underline{drinks} scotch, \[but \underline{not} Bill\].

(33)  
- a. Jean \underline{viendra à la fête} [et Marie \underline{aussi}].
  "Jean will come to the party and Marie too."
- b. Jean \underline{n’est pas venu à la fête} [et Marie \underline{non plus}].
  "Jean will not come to the party and Marie neither."
- c. Jean \underline{est venu hier} [mais \underline{pas} Marie].
  "Jean came yesterday, but not Mary."

- Besides the difference in the number of remnants, it is quite difficult to distinguish them in some languages (such as English), since they seem to have the same syntactic behaviour.
Same behaviour wrt embedding

- Unacceptability if only the target is embedded. According to Johnson (2009, 2014), there would be a very strong syntactic constraint on gapping (and a diagnostic of this elliptical construction), called ‘the no embedding constraint’, which does not apply in other cases, e.g. VPE (36).


(34) a. *Jane loves to study rocks, and John says that geography too. (Lobeck 1995)
b. ??/*John drinks scotch, and I think that Bill too.

(35) a. *Alfonso stole the emeralds, and I think that Mugsy the pearls. (Hankamer 1979)
b. *Some had eaten mussels and she claims that others shrimp. (Johnson 2009)
c. *Bill went to Paris and I think that John to Rome.
d. Bill went to Paris and I think that John went to Rome.

(36) The adults had eaten mussels, and she claims that the children did too. (Farudi 2013)
A not too strong constraint!

- Nevertheless, recent work (Weir 2014) shows that English can embed gapping too in some very restricted cases:
  - Absence of complementizer.
  - Gradient grammaticality related to the type of predicate (non-factive, semi-factive, true factive).

(37) John _ate_ oysters... (Weir 2014)

a. and I {think | believe | hope | suspect | was told | imagine} Mary swordfish.

b. and I {found out | remember | deny | know} Mary swordfish.

c. and I {am proud | am angry | am surprised} Mary swordfish.

- Need of experimental work to check these intuitions, since this is a debated issue (contradictory data (37)–(38)). Variables: with and without complementizer, semantic type of the verb, type of ellipsis.

(38) Who married who?

a. I think Sam, Mittie and Liz, Vinnie.

Cross-linguistic variety wrt embedding

- Embedding, in general, a fragment under a complementizer results in ungrammaticality in some languages (e.g. English, Dutch and German), but not in others, e.g. Spanish, Polish or Hungarian (cf. Stainton 2006, Vicente 2013, Temmerman 2013, Weir 2014, Fernández-Sánchez 2016, etc).

- As for Gapping, the only reported counter-example to Johnson’s No Embedding Constraint comes from Farsi (the variety of Persian spoken in Iran): Farudi (2013) observes that gaps are possible under a wide range of embedding verbs and argues that the subordinating heads are not parenthetical, but syntactically integrated.

- Inside Romance, we observe a non-unitary behaviour of Stripping and Gapping within a same language and across languages.
  - French allows embedded targets with Stripping, but not with Gapping.
  - Romanian and Spanish allow embedded targets with both constructions.
French data

- Embedded stripped clauses are acceptable, unlike embedded gapped clauses.
- Note that only stripped clauses with an adverbial particle such as *aussi* 'too' can be embedded; embedding fragments in general is not allowed in French (40), like in English (41).

(39)  
   a. Marie _viendra à la fête_ et _elle m’a dit que_ son mari aussi.  
       'Marie will come to the party and she told me that her husband will come too.'
   b. Marie _est partie_ et _je pense que_ Paul aussi.  
       'Marie left and I think that Paul too.'

(40)  
   – Qui vient? 'Who is coming?'
   a. – _*Je crois que_ Jean. 'I think that John.'
   b. – _Je crois_ Jean. / _Jean, je crois_. 'I think John.'

(41)  
   – Who left?
   a. – _*I think that_ John.
   b. – _I think_ John.
Gapped clauses cannot be embedded in French, regardless of the presence/absence of the particle *aussi* 'too' or the semantic type of the verb.

(42)  

a. *Paul a mangé une pomme et on m’a dit que Marie une orange.
   'Paul ate an apple and one told me that Marie (ate) an orange.'

b. *Paul a mangé une pomme et je crois que Marie une orange.
   'Paul ate an apple and one told me that Marie (ate) an orange.'
Stripping and fragments in general can be easily embedded, cf. Romanian data below.

(43) a. Eu sunt bine și sper că și tu.
    ‘I am fine and I hope that you too.’

b. Imi plac mult florile, și văd că și ție.
    ‘I like much the flowers, and I see that you too.’

(44) a. – Cine vine mâine? – Cred că Ion.
    ‘– Who will come tomorrow? – I think that John.’

b. – Cine a citit toate cărțile? – Bănuiesc că nimeni.
    ‘– Who read all the books? – I guess that nobody.’
Previous work on Spanish fragments (de Cuba & MacDonald 2013, Fernández-Sánchez 2016) insists on the crosslinguistic relevance of the semantic distinction (factive vs. non-factive predicates, cf. Hooper 1974) as a very strong constraint for fragment embedding: fragments can be embedded by non-factive predicates, but not by factives.

(45) − Quién robó las joyas? ’Who stole the jewels?’
   a. − {Creo | supongo | me imagino | pienso} que tu hijo.
      ’I {think | suppose | imagine} that your son.’
   b. − *{Lamento | sé | me sorprende | me desagrada} que tu hijo.
      ’I {regret | know | am surprised | dislike} that your son.’

Our attested data show that this constraint is too strong: in both Romance languages, fragments and gapping can be embedded under non-factive predicates (think, suppose, imagine, etc.), but also under semi-factive ones (find out, know, see, etc.), being excluded with factive predicates (regret, forget, amuse, etc.).
Parallelism constraints Evidence for a fragment-based analysis

Spanish data:

(46) a. **Creo que** nada más. (weak non-factive)
   ‘I think that nothing more.’

b. Le preguntaron dónde la había aprendido y **respondió que** en la calle. (strong non-factive)
   ‘They asked him where he had learnt it and he said that in the street.’

c. **Ya sabes que** yo un poco tímida al principio. (semi-factive)
   ‘You know that me a bit shy in the beginning.’

Romanian data:

(47) a. A luat Ion examenul și **cred că** și Maria.
   ‘Ion passed his examen and I think that Maria too.’

b. Eu sunt bine și **sper că** și tu.
   ‘I am fine and I hope that you too.’

c. Imi plac mult florile, și **văd că** și ție.
   ‘I like much the flowers, and I see that you too.’
Embedded Gapping in Romanian and Spanish

- Attested data from both Romance languages show that ‘the no embedding constraint’ is too strong crosslinguistically.

(48) Pedro le regaló flores a María y creo que Alicia unos libros.

‘Pedro offered flowers to Maria and I think Alicia some books.’

(49) Nu eu îl urăsc pe el, ci cred că el pe mine.

‘I don’t hate him, but I think he hates me.’
CORLEC Spanish Corpus (Corpus Oral de Referencia del Español Contemporáneo, Marcos-Marín 1992) furnishes data where gapping appears indeed in embedding contexts.

(50)  

a. Ella se lo va a comer todo pero me parece que yo solo un poco. (CONV 012A)  
‘She is going to eat everything but I think I only a bit.’

b. Pero el chico la ama y dicen que ella a él. (CONV 033A)  
‘But the boy loves her and they say she him.’

c. Luisa ha estado en ese club muchas veces pero por supuesto que yo nunca.  
‘Luisa has been in that club many times but of course me never.’
Both Stripping and Gapping allow embedding with relatively same kinds of semantic predicates in both Romanian and Spanish.

If we add to this semantic constraint the high frequency of the 1st person corroborated with the high frequency of epistemic predicates, we obtain a robust generalization concerning the pragmatic effect of the embedding construction: the speaker may take on some degree of commitment to the truth of the embedded fragment, leaving room for doubt and uncertainty.

We observe however that embedded gapping is sensitive to more criteria than embedded fragments in general (high frequency of true/false and epistemic predicates, symmetrical/reciprocal relations or reinforced semantic contrast), which is expected, since gapping, unlike other fragments, imposes very strong semantic and discursive constraints (Hartmann 2000, Kehler 2002).
Preference for symmetrical relations and reinforced semantic contrast, cf. Romanian data:

(51) a. Nu eu îl urăsc pe el, ci crez că el ne găsesc pe mine.
   ‘I don’t hate him, but I think he hates me.’

b. Ion o iubește pe Ana și văd că și ea pe el.
   ‘Ion loves Ana and I see that her him too.’

c. Ion este îndrăgostit de Ana, nu știu însă dacă și ea de el.
   ‘Ion is in love with Ana, but I don’t know if she him too.’
These new data contradict the common assumption that there is a universal ban on embedding and support a continuum analysis for embedding Stripping and Gapping crosslinguistically.

Both constructions do not have the same behaviour wrt embedding across languages and even within a same language.

No universal embedding constraint on Gapping! Typologically, there are three kinds of languages wrt embedded Gapping:

- Languages where embedded gaps are impossible (English, French).
- Languages where embedded gaps are possible only with some kinds of predicates (Romanian and Spanish).
- Languages where embedded gaps are possible under a wide range of non-parenthetical embedding verbs (Farsi, cf. Farudi 2013).

However, need of experimental evidence in order to check the semantic classes of verbs and the appropriate discursive context facilitating embedding.
We analyse these embedded fragments as a kind of unorthodox embedding:

- syntactically, a fragmentary clause is embedded under a main clause, but
- semantically, this main clause has a parenthetical use (i.e. it is the content of the embedded clause that is the main content of the utterance, cf. Hooper 1974) and
- pragmatically, the embedding predicate in the main clause has a discursive function, i.e. an evidential marker (indicating the speaker’s grounds for asserting the complement, cf. Simons 2007).

This unorthodox embedding construction recalls the complex interaction of multiple factors (syntactic, semantic and pragmatic) assumed by Green (1976: 382) to explain the embeddability of so-called Main Clause Phenomena:

"A number of syntactic constructions claimed by linguists to be restricted to main clauses are shown to occur, in fact, in a variety of subordinate clause types, but only under certain mysterious conditions – basically, when the speaker desires to be understood as committed to the truth of the subordinate clause."