32. Jahrestagung der DGfS, AG 6 "Deconstruction and Reconstruction of Portmanteau Morphemes" Humboldt University, Berlin, 24-25 February 2010

## French portmanteaus as simplex elements

Olivier Bonami<br>Université Paris IV<br>olivier.bonami@paris-sorbonne.fr

Jesse Tseng
University of Toulouse
Jesse.Tseng@univ-tlse2.fr

## 1 Introduction

4 portmanteau forms in modern French corresponding to [preposition + definite article] sequences:
(1)

|  | sing. |  |  | $\begin{gathered} \text { plur. } \\ \text { les }[\mathrm{le}(\mathrm{z})] \end{gathered}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  | le [lə] | la [la] | l' [1] |  |  |
| de [də] | du [dy] | de la | de l' | des [de(z)] | 'of the, from the, . |
| à [a] | au [o] | à la | à l' | aux $[\mathrm{o}(\mathrm{z})]$ | 'at the, to the, .. |
| en [ $\tilde{\mathrm{a}}(\mathrm{n})$ ] | (au [o]/' en le ) | en la | en l' | (ès [ ss ] / ' $\mathrm{en} \mathrm{les)}$ | 'in the, ...' |

(2)
a. Spanish: del $(=\mathrm{de}+\mathrm{el})$, $\mathbf{a l}(=\mathrm{a}+\mathrm{el})$
b. Italian: del $(=\mathrm{di}+\mathrm{il})$, della $(=\mathrm{di}+\mathrm{la})$, nel $(=\mathrm{in}+\mathrm{il})$, sul $(=\mathrm{su}+\mathrm{il}), \ldots$
c. Portuguese: do $(=d e+o)$, das $(=d e+a s)$, num $(=e m+u m), \ldots$
d. German: vom $(=$ von + dem $)$, zur $(=z u+$ der $)$, fürs $(=$ für + das $), \ldots$

- First issue: status of contracted form. We consider three possibilities: ${ }^{1}$
(3)
a. Phonological contraction

b. Multidomination
c. Simplex form



$$
\mathrm{al} \rightarrow \mathrm{o} / \ldots \mathrm{N}^{\prime}
$$

- We defend a simplex form analysis.
- Second issue: blocking of non-contracted forms (*à le garçon). Two strategies:
(i) Competition analysis: the non-contracted form is generated by the grammar but loses a competition with the contracted form.
(ii) Direct syntactic analysis: the grammar does not generate the non-contracted form.
- On the basis of coordination data we argue that (ii) is the better solution.


## 2 The data

### 2.1 Historical development of the forms

- etymological analysis: DE/A(D) + forms of ILLE (demonstrative pronoun $\leadsto$ def. article)

[^0]- but: portmanteau forms already obligatory in the earliest Old French texts ${ }^{2}$
(4) conservative forms in Vie de Saint Léger (10th cent.)
a. Del corps asaz l'avez audit et dels flaiels que grand sustint of-the body enough it-have heard and of-the afflictions that grandly sustained You have heard enough about the body and about the afflictions that he sustained with grandeur
b. al rei lo duistrent soi parent [...] als altres sanz en vai en cel to-the king him led his parents to-the other saints EN goes in heaven His parents took him to the king. He goes to the other saints in heaven.
- underlying forces leading to contraction (and elision)
- unstressed words beginning/ending in vowel
- proclitic nature of prepositions $\grave{a}$ and $d e$
- definite article forms < Latin ILLE/ILLA (proclitic and/or enclitic)
- further phonetic evolution can be explained, but can hardly be considered "regular"
(5) vocalization of final [1] in singular:
a. $[$ de le] $>$ del $>$ deu $(>$ dou) $>$ du
b. $[\mathrm{a}+\mathrm{le}]>\mathrm{al}>\mathbf{a u}$
(6) simplification of consonant cluster in plural:
a. $[$ de les] $>$ dels $>$ des
b. [a les] $>$ als $>$ as, later replaced by analogical aus $(=a u+s)$
- do not correspond to morphophonological processes productive in modern French
- [de le] must contract to du, [à les] to aux, etc., but no synchronic motivation available for the phonological result of this contraction
- irregular evolution of [en le] > au, [en les] > ès; forms no longer treated productively as contractions involving en


### 2.2 Contexts of (non)-contraction

### 2.2.1 Pronominal le, les

No contraction with $d e$ and $\grave{a}$ as infinitival markers followed by the accusative clitic pronouns le/les
(7) Nous avons peur du conseiller $_{N} /$ de le conseiller $_{V}$.
we have fear of.the adviser / of him advise
'We are afraid of the adviser.' vs 'We are afraid of advising him.'
(8) possible historically and dialectally ${ }^{3}$
a. et si s'acorderent al faire (Villehardouin, early 13th cent.)
and thus REFL-agreed to-it do
'And they agreed to do it.'
b. \% Jsuis ben content des voir (colloq. Quebec French)

I-am well happy of-them see
'I am very happy to see them.'
c. cf. modern standard French: ils s'accordèrent à le faire, content de les voir

[^1]
### 2.2.2 Atypical NPs in le

- No contraction with nouns beginning accidentally with [le] or [le]
(9) a. beaucoup de levain [ləv $\tilde{\varepsilon}]$ 'a lot of bread starter’ / beaucoup *duvain [dyvẽ] beaucoup *de le vin [ləvẽ]/ beaucoup du vin [dyvz̃] 'a lot of the wine'
b. piège à lémurs [lemys] 'lemur trap' / *aux-murs [отук] tarte *à les mûres [lemyь] / aux mûres [отуs] 'blackberry pie'
- With names:
(10) Person names
a. Surnames having the form of definite NPs never give rise to contraction.
i. Le Nôtre, Leclerc ( $\approx$ nôtre, clerc); Le Floch, Le Pen (*floch, *pen)
b. Nicknames do give rise to contraction.
i. "le Pirate", "le Chacal" ("the Jackal")
c. Interesting case: Le Corbusier: nickname based on an imaginary noun. Speakers hesitate between contraction (du Corbusier) and non-contraction (de Le Corbusier).
(11) Placenames
a. city names: Le Havre, les Lilas, Le Mans, Les Aubrais

Contraction is standard, but non-contraction is commonly heard, e.g.
i. Nous entrons en gare de Les Aubrais

We enter in station of Les Aubrais 'We are about to arrive at Les Aubrais station.'
b. Country, state or district names: le Maroc, le Maryland, le Languedoc Contraction is systematic

- Titles
(12) The title is a simple NP: Le Monde (newspaper), Les mouches (play by J.-P. Sartre) Contraction is systematic, but felt as somewhat strange. Much wordplay with non-contraction.
(13) The title is a coordinated NP: Le rouge et le noir (1830 novel by Stendahl), Le chagrin et la pitié (film by Max Ophüls), Le masque et la plume (radio show)
Both contraction and non-contraction are found, both are felt as bizarre.
a. J'ai lu hier soir, dans mon lit, le premier volume de Le rouge et le noir, de Stendhal (Gustave Flaubert in a letter to A. Le Poittevin, August 1845) "Last night in my bed I read the first volume of Le rouge et le noir, by Stendahl."
b. La chartreuse de Parme, par l'auteur du Rouge et le Noir (Théodore Muret, title of an article in La Quotidienne, July 24, 1839)
(14) The title is a non-NP starting with le or les: Le Roi se meurt (play by E. Ionesco), Les Héros sont fatigués (movie by Y. Ciampi)
Non-contraction is clearly preferred, although it is felt as problematic


### 2.2.3 Across the edge of a coordination

- When $\grave{a}$ and $d e$ take NP complements, they can have wide scope over a coordination of NPs, although repeating the preposition is more frequent.
(15) a. UroGène ( 55 salariés), spécialisée dans les cancers de la prostate et de la vessie, a ainsi racheté Chrysalon [...]. (Le Monde, August 2, 2003)
"UroGène, specializing in cancers of the prostate and of the bladder, acquired Chrysalon"
b. Alain Chergui a été blessé à la tête et à la poitrine par un éclat de métal. (Le Monde, August 23, 2003)
"A. Chergui was injured in the head and in the chest by a piece of metal."
(16) a. Ensuite pour la vie économique de la ville et la région, ce serait un tribut très lourd à payer [...] (Le Monde, July 7, 2003)
"Afterwards, for the economic life of the city and the region, it would be a very heavy price to pay"
b. La fédération marocaine a saisi, mercredi 31 janvier, son homologue française sur les cas de Jalil Narjissi et Hamid Arif [...] (Le Monde, February 2, 2007)
"The Moroccan federation formally contacted its French counterpart on Wednesday, January 31, with regard to the cases of J. Narjissi and H. Arif"
- Neither contraction nor non-contraction is possible with le or les in the initial conjunct.
(17) a. Le code de la famille nouvelle version reconnaît la responsabilité partagée du père et de la mère au sein de la famille. (Le Monde, February 25, 2005)
"The new version of the Family Code recognizes the shared responsibility of the father and of the mother within the family."
b. * de le père et la mère
c. * du père et la mère
- $\grave{A}$ and de cannot precede a coordination where a non-initial conjunct starts with le or les.
(18) a. Sur l'écran s'inscrit tout un monde de solitude autour de la mère et du fils. (Le Monde, April 4, 2004)
"Onscreen, a whole world of solitude is created around the mother and the son"
b. * de la mère et le fils
- Other languages (e.g. Spanish and German) use different strategies.
(19) portmanteau in initial conjunct without repetition of preposition
a. El Extraño caso del hombre y la bestia (Argentinian film)
"The strange case of the man and the beast"
b. dass sowohl bei der Fertigung, als auch beim Transport und der Anwendung keine Gefahr für den Menschen besteht ${ }^{4}$
"that during preparation, as well as during transport and use, there is no danger to people"
(20) contracting form of article in non-initial conjunct without repetition of preposition
a. la última campaña de la marca Adidas, que cuenta con figuras de la música y el deporte ${ }^{5}$ "the latest campaign by Adidas, which relies on stars from music and sports"
b. weil Irrtümer bei der Textabfassung und dem Druck nicht ausgeschlossen werden können ${ }^{6}$ "because mistakes during text composition and printing cannot be ruled out"


## 3 Generating contracted forms

### 3.1 Against a contraction analysis

- No plausible synchronic phonological rule predicts the form of portmanteaus

[^2]- The use of portmanteau forms is sensitive to fine details of syntactic analysis
- The very formulation of the rule is problematic
- We need to know the syntactic structure of the input

$$
\begin{equation*}
\mathrm{a}_{1} \# \mathrm{l}_{2} \rightarrow \mathrm{o} /\left[\mathrm{PP} \__{2}\left[\mathrm{NP} \__{2} \mathrm{~N}^{\prime}\right]\right] \tag{21}
\end{equation*}
$$

- but we do not know where to put the output in the syntactic structure, without modifying that structure. But if we modify it then this is syntax, not phonology.


### 3.2 Against multidomination

- Multidomination analyses have been proposed as a solution for syntactic portmanteaus by e.g. Blevins (1990); Wescoat (2002); Payne et al. (2007). Here we rely on Wescoat's proposal, but our criticism extends to all cases.
- Wescoat (2002): terminals may correspond to sequences of adjacent preterminals
(22) Toy lexicon: de: P le: Det la: Det du: P Det fille: N garçon: N
a.

b.

c.

- This simple analysis needs to be supplemented by some theory of blocking, because the grammar generates non-contracted forms (23c).
(24) Tentative blocking principle for Wescoat (2002)

Two trees are said to be equivalent if they are identical up to nonterminals. When two equivalent trees are such that their respective terminals express the same morphosyntactic information and the same semantic content, only the tree which minimizes the number of terminals is grammatical.

- Problems:
- As it stands the principle is rather vague. How does one check that two trees express the same information? Should word order be taken into account?
- Without additional constraints, this approach to blocking predicts that either (25a) or (25b) should be possible, contrary to fact.
(25)

a.


b.

- Conclusion: multidomination does not provide an account of French portmanteaus, even supplemented with a theory of blocking.


### 3.3 Portmanteaus as simplex forms

- Contracted forms are special prepositions which:
- subcategorize for an $\mathrm{N}^{\prime}$ rather than an NP (26a)
- incorporate the syntactic and semantic features of the definite article

See section 5 for explicit lexical entries

- The exclusion of contraction at the edge of coordinations reduces to a subcategorization violation (26b)
(26)
a.

b.

- Advantages:
- The analysis of French portmanteaus reduces to familiar phenomena
- Accounts straightforwardly for the impossibility of contraction at the edge of a coordination
- But we still need a way of blocking non-contracted forms, including in cases of coordination


## 4 A competition-based account

- Remember: we need a way of excluding
(27) a. *de le père
b. *de le père et le fils
c. *de la mère et le fils
(blocking in simple phrases)
(blocking by initial conjuncts)
(blocking by non-initial conjuncts)
- We first attempt a competition-based approach:
- Generate non-contracted forms
- Set up a competition (OT-style) with contracted forms, and exclude the losers.


### 4.1 Two functional constraints

- We propose a constraint: Be brief!
- Clear functional motivation. Remember that this is a facet of Grice's Maxim of Manner.
- Can be implemented as either counting (overt) terminals or syllables.
(28)

| of ( $\sigma x$.father $(x)$ ) | BE BRIEF! |
| :---: | :---: |
|  | * |
| 17 |  |

- In cases of coordination, the candidate set must include different ways of factoring constituents.
(29) a. J' ai parlé [de la mère et de la fille].

I have spoken of the mother and of the daughter
"I spoke about the mother and the daughter."
b. J' ai parlé [de la mère et la fille].

I have spoken of the mother and the daughter

- The fact that both examples in (29) are grammatical shows that BE bRIEF! must be counterbalanced by another constraint.
- We propose a constraint Be redundant!
- Maximize the number of exponents of every morphosyntactic, syntactic, or semantic feature.
- Functional motivation: help the addressee get the correct signal over a noisy channel.
- Can be seen as a corollary of Grice's Avoid ambiguity.
- Because they must balance each other, Be brief! and Be redundant! are unordered.
(30)



### 4.2 Portmanteaus and coordination

- If we have the option of using portmanteaus, the equilibrium is lost, because a portmanteau is more informative with fewer segments.
(31)

|  | of $(\sigma x$.father $(x) \oplus \sigma x$.mother $(x))$ | Be brief! | Be Redundant! |
| :---: | :---: | :---: | :---: |
|  |  | * | * |
|  |  | ** |  |
| Ler |  | * |  |

- Now we also have a way of accounting for the blocking at a distance of non-contracted forms.
(32)

|  | of $(\sigma x$.mother $(x) \oplus \sigma x . \operatorname{son}(x))$ | BE BRIEF! | Be Redundant! |
| :---: | :---: | :---: | :---: |
|  |  |  | * |
|  |  | * |  |
| (1) |  |  |  |

- Despite its appeal, there are a number of problems with this analysis.
- Situations such as those in Spanish and German (cf. (19-20)) are predicted not to occur: Be redundant! needs to be dropped, but this amounts to a general ban on repetition in coordination.
- Formulation of Be brief!:
* If we count phonological length, then all phonological reduction phenomena should be obligatory, contrary to fact.
* Even if we count overt terminals, we predict that e.g. contracted negative forms in English should be obligatory.
(33) But I'm definitely better than I used to be, and improving all the time, because over and over and over I learn that I will not get what I do not ask for. And you won't either. ${ }^{7}$
- Formulation of Be redundant!:
* Evasive formally. How does one count features? What counts as an exponent?
* Be redundant! excludes situations where fewer features may be expressed despite identical length, contrary to fact.
a. Je ne suis pas content ou mécontent.

I NEG am NEG happy or unhappy
'I am neither happy nor unhappy.'
b. Je ne suis pas content ni mécontent.

I NEG am NEG happy or.NEG unhappy 'I am neither happy nor unhappy.'

## Conclusion:

- Functional considerations may contribute to shaping the grammar over time, and help us to understand some aspects of its synchronic organization.
- But grammatical constraints are not simply hard-wired functional constraints (and often have no functional motivation whatsoever).


## 5 HPSG account

We propose a lexical approach to the analysis of the portmanteau forms themselves, and a syntactic approach to the coordination facts.

### 5.1 Lexical entries of simplex forms

(35)
a. $a u$

b．$d u$

$$
\left[\begin{array}{ll}
\text { PHON } & \langle\mathrm{dy}\rangle \\
\text { HEAD } & {\left[\begin{array}{ll}
\text { prep } \\
\text { PFORM } & d e
\end{array}\right]} \\
\text { VAL } & {\left[\begin{array}{ll}
\text { COMPS } & \left\langle\mathbf{N}^{\prime}{ }_{m s g}[- \text { LTRIG }]\right\rangle
\end{array}\right]}
\end{array}\right]
$$

（36）plural forms
a． $a u x$
$\left[\begin{array}{ll}\text { PHON } & \langle\mathrm{o}(\mathrm{z})\rangle \\ \text { HEAD } & {\left[\begin{array}{ll}\text { prep } & \\ \text { PFORM } & \grave{a}\end{array}\right]} \\ \text { VAL } & {\left[\begin{array}{ll}\text { COMPS } & \left\langle\mathbf{N}^{\prime}{ }_{\text {plur }}\right\rangle\end{array}\right]}\end{array}\right]$
b．
des
$\left[\begin{array}{ll}\text { PHON } & \langle\operatorname{de}(\mathrm{z})\rangle \\ \text { HEAD } & {\left[\begin{array}{ll}\text { prep } \\ \text { PFORM } & d e\end{array}\right]} \\ \text { VAL } & {\left[\begin{array}{ll}\text { COMPS } & \left\langle\mathbf{N}^{\prime}{ }^{\text {plur }}\right\rangle\end{array}\right]}\end{array}\right]$
－Combination of the categorial properties of the prepositions $\grave{a}$ and $d e$ and the syntactic effects of the definite article（i．e．the complement must be an $\mathrm{N}^{\prime}$ with the appropriate agreement features）
－Combination of the semantic content of the preposition and the definite article（not indicated above），e．g．for locative $\grave{a}$ ：［CONTENT 〈loc－at－rel，def－rel 〉］
－PFORM value used for syntactic selection of prepositions： appartenir à Paul，au roi，aux Allemands＇belong to Paul，to the king，to the Germans＇
－［ $\pm$ LTRIG $]$ is determined by the leftmost word of the phrase，and correlated with its first segment （vowel vs consonant）：

－［－LTRIG］：au méchant ennemi＇to the wicked enemy＇，du directeur＇from the director＇
－The liaison alternation in the plural is also determined by the LTRIG value of the $\mathrm{N}^{\prime}$（Bonami et al．，2004）：
－［＋LTRIG］：aux énnemis［ozenmi］，des anciens directeurs［dezãsjẽdiьをktøь］
－［－LTRIG］：aux méchants ennemis［omefãzenmi］，des directeurs［dedibદktœв］
（37）
a．$\grave{a}$

b. $d e\left(d^{\prime}\right)$
$\left[\begin{array}{lll}\text { Phonology } & \langle\mathrm{d}(\boldsymbol{\theta})\rangle \\ \text { Head } & \left.\begin{array}{ll}\text { prep } \\ \text { PForm } & d e\end{array}\right] \\ \text { valence } & {\left[\begin{array}{ll}\text { Complements } & \langle\text { NP }[- \text { Le }]\rangle\end{array}\right]}\end{array}\right]$

- The $d e \sim d^{\prime}$ alternation is determined by the ltrig value of the NP.
- [-LTRig]: de la mère [dəlamعь], de Paul [dəpıl]
- [+Ltrig]: d'une mère [dynmes], d'Albert [dalbes]

We introduce another feature $[ \pm \mathrm{LE}]$ that propagates along the left edge of simple phrases.

- The contracting forms of the definite article le and les have the value [+LE].
- All other words, including $l a$ and $l$ ', have the value [-LE].
- The specification [-LE] in (37) prevents $\grave{a}$ and de from combining with a complement beginning with $l e$ or les.
- accounts for the blocking effect: *de les amis, *à le roi
- determined by the leftmost word of the NP: de tous les amis 'to all the friends', à feu le roi 'to the late king'
(38) Propagation of $\pm$ LE in simple phrases
a.

b.



### 5.2 Portmanteaus in coordinated phrases

The simplex analysis and the [-LE] constraint allow us to account for the ungrammaticality of both contraction and non-contraction in initial conjuncts, cf. (17), (25)
(39) Blocking by [+LE] initial conjunct
a.

b.


- In (39a), the NP conjunct is incompatible with the subcategorization requirements of of au.
- In (39b), the coordinated phrase has the feature [+LE], and thus cannot combine with $\grave{a}$.
- At first sight, this looks like propagation of [+LE] along the left edge of the phrase, cf. (38).
- In fact, in coordinated phrases, the value of LE is determined as a function of the LE value of all conjuncts.
A single $[+\mathrm{LE}]$ conjunct, in any position, causes the entire coordinated phrase to be [ +LE ], and thus unacceptable as a complement of $\grave{a}$ or $d e$.
- This feature resolution analysis allows us account for blocking by non-initial conjuncts (40a) and initial conjuncts (39b) in the same way.
(40) Resolution of $\pm$ LE in coordinated phrases
a.

b.


More familiar examples of feature resolution in coordinated phrases, where one value takes precedence in unlike combinations:

- 1st person \& 2nd/3rd person $\rightarrow$ 1st person
- 2 nd person $\& 3$ rd person $\rightarrow 2$ nd person
- feminine \& masculine $\rightarrow$ masculine
- Resolution is a syntactic rule, but distance can interfere (performance effect):
(41) a. [Ma mère, ma belle-sœur, et ma nièce] sont très heureuses.
$(\mathrm{F}+\mathrm{F}+\mathrm{F} \rightarrow \mathrm{F})$
b. [Ma mère, ma belle-sœur, et mon frère] sont très heureux.
$(\mathrm{F}+\mathrm{F}+\underline{\mathrm{M}} \rightarrow \underline{\mathrm{M}})$
c. [Mon frère, ma belle-sœur, ma nièce et ma mère] sont ?heureux/?? heureuses. my brother my sister-in-law my niece and my mother are happy-masc/happy-fem
(42)
a. J'ai parlé *[à Paul et le chef] $([+\mathrm{LE}])$ / [à Paul et au chef].
b. J'ai parlé [à Paul, Henri, nos collègues bulgares, tous les étudiants et le chef]. I spoke to Paul Henri our Bulgarian colleagues all the students and the boss


## References

Abeillé, A., Bonami, O., Godard, G., and Tseng, J. (2006). The syntax of French $\grave{a}$ and de: an HPSG analysis. In P. Saint-Dizier (ed.), Syntax and semantics of prepositions, p. 147-162. Berlin : Springer.
Blevins, J. P. (1990). Syntactic complexity: evidence for discontinuity and multidomination. Ph.D. thesis, UMass Amherst.
Bonami, O., Boyé, G., and Tseng, J. (2004). An integrated approach to French liaison. Proceedings of Formal Grammar 2004, pages 29-45. Loria, Nancy.
Grevisse, M. and Goosse, A. (2008). Le Bon Usage. De Boeck Duculot, Brussels, 14th edition.
Miller, P. (1991). Clitics and constituents in Phrase-structure grammar. Ph.D. thesis, Utrecht U.
Nyrop, K. (1908). Grammaire historique de la langue française. Gyldendal, Copenhague.
Payne, J., Huddleston, R., and Pullum, G. K. (2007). Fusion of functions: The syntax of once, twice, and thrice. Journal of Linguistics, 43, 565-603.
Wescoat, M. T. (2002). On lexical sharing. Ph.D. thesis, Stanford University.


[^0]:    ${ }^{1}$ We do not assume a DP analysis, but this is orthogonal to the current issue. Substitute DP for NP and NP for $\mathrm{N}^{\prime}$ if this is your taste.

[^1]:    ${ }^{2}$ For rare examples of non-contraction, see Nyrop (1908, §503).
    ${ }^{3}$ Ex. (8b) cited by Grevisse and Goosse (2008, §659, f).

[^2]:    ${ }^{4}$ http://www.industrystock.de/html/Brandschottung/product-result-de-52144-0.html
    ${ }^{5}$ http://www.malianteo.com/general/calle-13-aparece-en-anuncio-de-adidas-de-star-wars/
    ${ }^{6}$ http://www.it-recht-kanzlei.de/irrtuemer-vorbehalten-agb.html

