



# Phrasal affixes and French morphosyntax

Jesse Tseng

## ► To cite this version:

Jesse Tseng. Phrasal affixes and French morphosyntax. Proceedings of Formal Grammar 2003 FGVienna, Gerhard Jäger, Paola Monachesi, Gerald Penn, Shuly Wintner, May 2003, Vienne, Autriche, pp.177-188. inria-00107734

**HAL Id: inria-00107734**

**<https://inria.hal.science/inria-00107734>**

Submitted on 19 Oct 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***Proceedings***

of

***Formal Grammar 2003***

***Gerhard Jäger, Paola Monachesi, Gerald Penn and Shuly Wintner  
(editors)***

## Chapter 14

---

# Phrasal affixes and French morphosyntax

JESSE TSENG

*MoDyCo (UMR 7114), Université Paris X – Nanterre*  
*E-mail: jtseng@linguist.juissieu.fr*

**ABSTRACT.** A number of “minor” elements in French occupy the boundary between syntax and morphology and call for a special analysis. This paper presents an HPSG formalization of a proposal for treating *à*, *de*, and the definite article *le* as phrasal affixes. The mixed status of these forms requires a two-level analysis, with morphophonological realization at the lexical level (using lexical prefixation rules) and syntactic/semantic interpretation in the syntax (using non-branching ID rules). The two levels are linked by EDGE feature percolation. The proposed treatment accounts for a range of grammatical phenomena that are highly problematic for traditional analyses in which *à*, *de*, and *le* are considered to be independent syntactic words.

Many languages have elements that are difficult to classify conclusively as words or bound morphemes, since they share some grammatical properties with both categories. The formal analysis of these elements is best approached within a theoretical framework in which the interface between morphology and syntax (as well as their interaction with phonology) can be represented. This paper presents one particularly problematic interface phenomenon in French—the case of phrasal affixes—and proposes a lexicalist analysis in HPSG.<sup>1</sup>

### 14.1 Phrasal affixation

It is useful to begin by mentioning the case of pronominal clitics in French.<sup>2</sup> On the one hand, clitics have the same syntactic and semantic functions as normal NP and PP arguments, and treating them as independent syntactic items is the simplest way to explain this. On the other hand, clitics do not appear in the same positions as full NPs and PPs, they form a phonological unit with the verb, and they exhibit paradigmatic and combinatorial idiosyncrasies that are more characteristic of bound morphemes. For example, the same clitic cannot occur twice on the same

---

<sup>1</sup>The proposals in this paper were developed as part of a French grammar implementation project within the research groups TaLaNA (UMR 8094) and the Laboratoire de Linguistique Formelle (UMR 7110) at Université Paris 7. The author wishes to acknowledge the support of these groups.

<sup>2</sup>Similar remarks apply to the other Romance languages (see Miller and Monachesi, in press for a descriptive survey).

verb (even with distinct grammatical functions), and the relative order of clitics is completely fixed (according to case, person, reflexivity, and prefix/suffix status). A number of analyses are possible for these exceptional elements; recently Miller and Sag (1997) have argued convincingly that French clitic pronouns are affixes, attached lexically to the verb, and not introduced by syntactic combination. They provide an HPSG account of clitic realization that handles both the morphophonological facts and the interaction with verbal argument structure.

This paper adopts many of the ideas of Miller and Sag, applying them to another set of elements in French: the definite article *le* and the prepositions or markers *de* and *à*. Miller (1992) demonstrates that the same criteria that decide in favor of affixal status for clitic pronouns also apply to these elements. Some of the relevant tests are illustrated here:

- (1) failure to coordinate with other elements
  - a. \**de* ou *à* Bruxelles, \**à* et après minuit
  - b. \**le* et la secrétaire(s), \**les* ou d'autres lettres
- (2) lack of wide scope over a coordinate structure
  - a. aimer [le poisson et \*(le) fromage]
  - b. appartenir [à l'Etat ou \*(à) l'Eglise]
  - c. vins [de pays et \*(de) table]
- (3) haplology effects
  - a. réfléchir [à des livres] vs. parler [de (\*des) livres]
  - b. le livre [le plus important] vs. le [(\*)le plus important] livre

As noted by an anonymous reviewer, and by Miller himself, the results of these tests are not always consistent. For example, coordination is possible in cases of metalinguistic number or gender uncertainty (*le ou la*, *le ou les*), and wide scope is possible in some situations.<sup>3</sup> It should be kept in mind, however, that even uncontroversial derivational affixes sometimes "fail" these tests (e.g., German *Vor- und Nachspeisen* 'appetizers and desserts'). It is important to consider the results observed for *à*, *de*, and *le* in comparison to other determiners and prepositions, like *plusieurs* or *avant*, which are not at all subject to the same constraints.

In addition to the strong indications provided by the tests above, the existence of the idiosyncratic (and obligatory) contracted forms *du*, *des*, *au*, and *aux* also motivates a morphological treatment of these elements. In a lexicalist framework like HPSG, these forms cannot be analyzed as the result of a combination of two syntactically independent elements. Finally, phonologically conditioned elision effects (involving the forms *d'* et *l'*) and obligatory liaison (with *les*, *des* et *aux*) indicate a close link with the following word, closer than that typically found between elements that are simply adjacent in the surface word order.

<sup>3</sup>And finally, the term "haplology" is most probably the wrong characterization for the phenomena illustrated in (3).

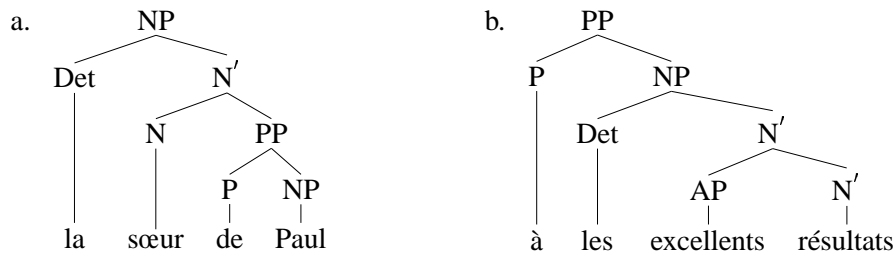


Figure 14.1: Traditional constituent structures

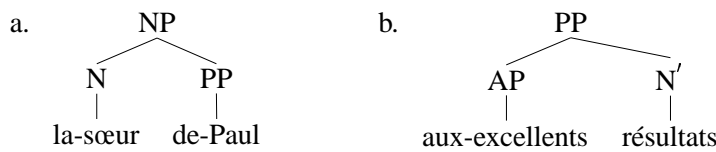


Figure 14.2: Phrasal prefixation

The grammatical evidence thus suggests that the traditional syntactic structures associated with these elements (see the examples in Figure 14.1) must be replaced by structures involving prefixes (Figure 14.2). We can see from these examples why *le*, *de*, and *à* are more difficult than clitic pronouns to analyze as affixes. Clitics are word affixes—i.e., they attach morphologically to a single word, and their syntactic and semantic effects are registered on that word. Furthermore, they are always arguments, syntactically and semantically, so they are not needed to select or introduce anything else. For example, *donner* is a ditransitive verb, expressing a three-place relation  $\text{give}(x,y,z)$ . But it can undergo clitic prefixation to become *me-les-donner*, and the effect of the two prefixes is to reduce the verb's valence to that of an intransitive verb, and to supply pronominal reference for two of the verb's semantic arguments. This is essentially the analysis formalized by Miller and Sag (1997).

In Figure 14.2, on the other hand, the prefixes must be *phrasal* affixes. This does not mean that they combine morphologically with a phrase; the principle of strict lexicalism requires affixation to take place at the lexical level, and so the host is still a single word. But the prefixes *le*, *de*, and *à* have syntactic and semantic scope over an entire phrase. In other words, they are functors, responsible for introducing a dependent phrase (an  $N'$  for *le*, an NP or VP for *à* and *de*). The problem is that this phrase is not (necessarily) identifiable at the point where the prefix is introduced. In some cases, as in Figure 14.2(a), this problem can be avoided: the NP “argument” of *de* happens to be a single word (*Paul*), and the lexical host of *la* happens to be the syntactic head (*sœur*) and can therefore provide a pointer to the syntactic and semantic features of its projection (at least in this example). But in general, the lexical host of the prefix does not carry enough information to allow

the syntactic and semantic functions of the prefix to be properly incorporated. This is the case in example Figure 14.2(b): the adjective does not encode the identity of the N' *excellents résultats* in any directly accessible way.

Given the very rich lexical descriptions used in HPSG, one must be careful not to overlook any potentially useful “buried” information. In fact, the HPSG analysis of modification (in which the modifier does contain a representation of the modified element in its MOD value) leaves room for an extremely inelegant work-around for the example in Figure 14.2(b), which could also be extended to modifier chains as in *les-très excellents résultats*. But cases involving a coordination of modifiers (e.g., *de-belles et grosses carottes*) seem to admit no lexical solution. Post-nominal modification (e.g., *l'argent que j'ai perdu*) also poses a problem, since under standard assumptions, a head does not encode the identity of its modifiers.

The cases of phrasal affixation in French can be understood from a diachronic perspective, assuming that the structures in Figure 14.1 were valid in earlier stages of the language. Over time, the elements in question lost their status as independent words (and their syntactic functions of determiner, preposition, complementizer). The change from syntactic word to morphological affix is consistent with the development of contracted forms and other irregularities, while the former syntactic function (specifier or head) accounts for the position of these elements at the left periphery of the phrase they scope over.

## 14.2 Edge features

According to the present proposal, in the current stage of French, there is a mismatch between the morphological scope of *le*, *de*, and *à* (a single word) and their syntactic and semantic scope (a phrase). The formal analysis of the elements, therefore, must similarly be split between two levels. The first step, prefixation, is handled by morphophonological rules in the lexicon, while the second step, interpretation, is handled by unary syntactic rules. The two levels are linked by information encoded in EDGE features, which propagate to the mother from the peripheral daughters in all branching syntactic combinations. The idea of EDGE features for determining affix or clitic placement was introduced by Klavans (1985). The HPSG analysis presented here is heavily influenced by the GPSG formalization of Miller (1992), and has been tested (in somewhat modified form) as part of a theory-driven French grammar implementation using the LKB platform for HPSG-style grammar development (Copestake, 2002; Tseng, 2003b, to appear).

Concretely, prefixation of *le*-, *de*-, and *à*- is assumed to apply quite freely to existing words. Multiple realizations of the same prefix are prohibited, however, and elision and liaison effects are built in to the rules as allomorphy phenomena. Also, prefixation of *de*- and *à*- is properly constrained so that *\*à/de+le/les* are blocked, in favor of *au*, *aux*, *du*, and *des*.

We have seen that at the point where prefixation is realized, the syntactic and semantic contribution of the prefix cannot be integrated, since identity of the

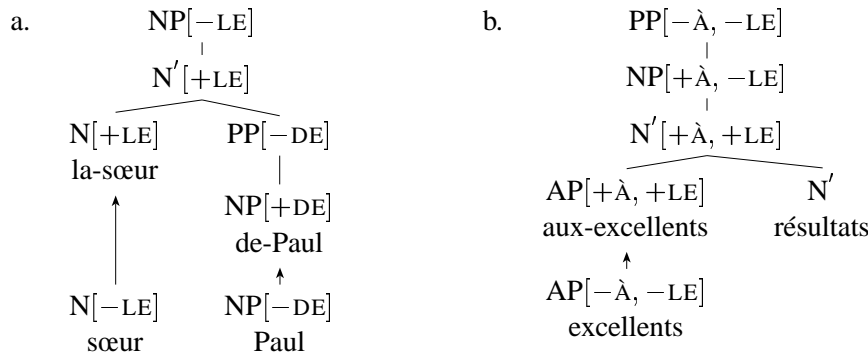


Figure 14.3: Phrasal prefixation with interpretation

phrasal argument cannot always be determined. In fact, prefixation has no immediate grammatical effect on the word involved. For example, *Paul* is a proper noun, and its prefixed forms *à-Paul* and *de-Paul* are also proper nouns. The adverb *très* is still an adverb after prefixation: *de-très*, *la-très*, *aux-très*, etc. But prefixation does add one or more positive EDGE specifications to the affected lexical item. The EDGE features needed for this analysis can be represented (for the moment) simply as boolean features  $[\pm A]$ ,  $[\pm DE]$  and  $[\pm LE]$ .<sup>4</sup> A rule for *de*-prefixation, for instance, takes a  $[-DE]$  word, prefixes *de-* (or *d'-*) to it, and switches its specification to  $[+DE]$ . A word prefixed with *du* has positive values for both DE and LE, and so on.

In the syntax, (left) EDGE specifications are copied from the left-most branch in every syntactic combination to the mother. In this way, syntactic rules have access to information regarding the prefixes on the left edge of every phrase. In particular, if a positive EDGE value is present on a phrase that the corresponding prefix could scope over, a unary syntactic rule can apply, modifying the phrase by incorporating the syntactic and semantic effects of the prefix. After the application of a prefix interpretation rule, the corresponding positive EDGE value is removed.

The two-step analysis of phrasal affixes is illustrated in Figure 14.3, an expanded version of Figure 14.2. The arrows indicate lexical rule application. The two derivations show the introduction of positive EDGE features by prefixation rules, and the percolation of EDGE values along the left periphery of phrases. Unary syntactic rules apply when phrases satisfying particular descriptions become available in the syntax. For example, any NP can serve as the argument of the prefixes *à-* or *de-*, and so any NP carrying the feature  $[+À]$  or  $[+DE]$ , indicating the presence of a prefix that has not yet been interpreted, can be transformed into a PP, with the positive EDGE feature “discharged.” Similarly, an N' with a non-negative value

<sup>4</sup>To be precise, in this paper we are dealing with EDGE | LEFT features, corresponding to the left periphery of the phrase, while EDGE | RIGHT features are needed for phrasal suffixes, like French *-ci* and *-là*. Left and right EDGE features interact in the analysis of consonant liaison in French (Tseng, 2003a).

for LE can become a full,  $[-LE]$  NP. A well-formed maximal projection should bear only negative EDGE values. Ungrammatical prefixes introduce non-negative specifications that cannot be discharged—for example, *\*le-Paul*, *\*à-excellent du-résultat*.

The structures in Figure 14.3 only show the syntactic effects of prefix interpretation (changes in syntactic category or bar level), but the unary interpretation rules also activate prefixes' semantic contribution (definiteness or indefiniteness, locative/directional or possessive meaning, etc.) Note also that the addition of the non-branching nodes reveals a layered syntactic analysis much more similar to the traditional conception in Figure 14.1 than first suggested by the structures given in Figure 14.2.

### 14.3 Formalization

In the remaining sections of this paper, elements of the formal HPSG analysis are presented in more detail, and the treatment of a number of more complex phenomena is discussed.

#### 14.3.1 Lexical rules

The morphophonological realization of phrasal affixes is handled by a set of lexical rules of type *word-to-word*. The significance of this is that *w2w* rules are assumed to apply optionally, since words are available for syntactic realization without any further lexical operations (as opposed to lexemes, which must undergo at least one *lexeme-to-word* lexical rule application).

Consider the following rule, responsible for the prefixation of *la-*:

$$(4) \quad \left[ \begin{array}{c} \text{word} \\ \text{PHON} \quad \boxed{1} \\ \text{EDGE} \quad \left[ \begin{array}{cc} \text{LIAIS} & - \\ \text{LE} & \text{none} \\ \text{À} & \text{none} \\ \text{DE} & \text{none} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{c} \text{word} \\ \text{PHON} \quad \text{la} + \boxed{1} \\ \text{EDGE} \quad \left[ \begin{array}{cc} \text{LIAIS} & - \\ \text{LE} & \left[ \begin{array}{c} \text{phr-aff} \\ \text{IN} \quad \text{N}' \quad [\text{fem}, \text{sing}] \\ \text{OUT} \quad \text{NP} \quad [\text{def}] \end{array} \right] \end{array} \right] \end{array} \right]$$

First of all, the feature  $[\pm\text{LIAIS}]$  is used to account for liaison and elision phenomena (Tseng, 2003a). For the most part,  $[-\text{LIAIS}]$  elements are consonant-initial words (and phrases) that fail to trigger liaison or elision to the left. The rule in (4) therefore selects a consonant-initial input word, and adds the prefix *la-* (so the result is again  $[-\text{LIAIS}]$ ).

The technical implementation of the prefix features À, DE, and LE is somewhat more complex than suggested in the previous section. The value of each feature is



either *none* or an object of type *phrasal-affix*. The type *phr-aff* introduces two *sign*-valued attributes IN and OUT that encode the potential grammatical effects of the prefix. In the *la-* example, the input must be [LE: *none*], in order to block iteration (*\*la-la-table*). The input is also specified as [DE: *none*] and [À: *none*], to prevent the ungrammatical prefix sequences *\*la-de-* and *\*la-à-* (this last case would also be excluded by LIAIS constraints).

In the output, the prefixed word has a “positive” LE specification. Informally, the IN value is a description of the sort of phrase the article *la* would combine with in a traditional analysis (cf. Figure 14.1), and the OUT value is a description of the phrase that would result from this combination. In the proposed phrasal affix analysis, there is no article *la*, but the prefix *la-* provides the instructions for generating the same result later on in the syntax. A priori, there is no link between the IN and OUT values associated with a given prefix, but we can assume informally that any features not explicitly mentioned in the OUT specification are shared with the IN value.<sup>5</sup>

Note that the rule in (4) overgenerates, since without additional constraints, nothing prevents the formation of prefixed forms like *\*la-Paul*, *\*la-chevaux*, *\*la-comment*. The LE *mid* IN value does specify the features *feminine* and *singular*, but the IN and OUT values are only activated in the syntax. The application of the prefixation rule could be limited to certain syntactic categories satisfying certain descriptions, but in any case, ungrammatical occurrences of the prefix will be rejected as uninterpretable by the syntax and will not give rise to syntactic over-generation.

The elided form *l’-* is handled by a separate, but similar rule, in which the input word is specified as [+LIAIS] (and the gender of the N’ in the output’s LE | IN value is left underspecified). The rules for the other forms *le-*, *les-* (with and without liaison), and for *de-*, *d’-*, and *à-*, follow the same model. Note that in general, each prefix will be associated with several distinct IN and OUT pairs, reflecting distinct syntactic and semantic functions, just as in traditional approaches, *à*, *de*, and *le* require multiple lexical entries.

The so-called “portmanteau” contractions (*au*, *aux*, *du*, *des*) are analyzed in two steps: prefixation of *le-* or *les-* followed by prefixation of *à-* or *de-*, but the output of the second step is irregular. The derivation of *aux-excellents* from Figure 14.3(b) is shown below:

(5)

$$\begin{array}{c}
 \left[ \begin{array}{l} \text{PHON} \text{ “excellents”} \\ \text{EDGE} \left[ \begin{array}{l} \text{LIAIS } + \\ \text{LE } \textit{none} \\ \text{À } \textit{none} \\ \text{DE } \textit{none} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{l} \text{PHON} \text{ “les-excellents”} \\ \text{EDGE} \left[ \begin{array}{l} \text{LIAIS } - \\ \text{LE } \boxed{1} \textit{phr-aff} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{l} \text{PHON} \text{ “aux-excellents”} \\ \text{EDGE} \left[ \begin{array}{l} \text{LIAIS } + \\ \text{LE } \boxed{1} \textit{phr-aff} \\ \text{À } \textit{phr-aff} \end{array} \right] \end{array} \right]
 \end{array}$$

<sup>5</sup> A more precise formulation of this idea is proposed in §14.3.3.

This analysis is motivated in part by the treatment of allomorphy in the lexical rule component of the LKB system. A more monotonic approach using a single rule could also be adopted. In either case, the final output crucially bears positive values for two EDGE features.

### 14.3.2 Propagation

A positive EDGE specification indicates the presence of a phrasal affix whose semantic and syntactic potential has not yet been realized. This realization can only happen when an appropriate phrase (matching the prefix's IN value) is constructed by the syntax. Until such a phrase becomes available, the EDGE specifications must remain visible at each successive level of syntactic structure. As we have seen, the propagation of (left) EDGE features is not necessarily head-driven; instead, it is systematically driven by the left-most daughter. Formally:

(6) EDGE Feature Principle

$$\left[ \begin{array}{c} \text{phrase} \\ \text{ARGS} \langle [ ] \rangle \oplus \text{nelist} \end{array} \right] \Rightarrow \left[ \begin{array}{c} \text{EDGE} [1] \\ \text{ARGS} \langle [ \text{EDGE} [1] ] \rangle \oplus \text{nelist} \left( \text{EDGE} \left[ \begin{array}{c} \text{LE} \text{ none} \\ \text{DE} \text{ none} \\ \text{À} \text{ none} \end{array} \right] \right) \end{array} \right]$$

The ARGS value lists all daughters, according to their surface order.<sup>6</sup> The first ARGS element thus corresponds to the left-most branch. The constraint above states that in any branching syntactic structure, the EDGE features of the left-peripheral daughter are shared with the dominating phrase. The other daughters are required to bear the value *none* for all prefix features.

The effects of this principle can be verified in the analyses shown in Figure 14.3. For example, it is the EDGE features of the adjective in Figure 14.3(b) that appear on the  $N'$ . The string *excellents aux-résultats*, with the prefix attached to the head N, would not allow the same analysis. In Figure 14.3(a), the positive value [+DE] of *de-Paul* must be discharged before this prefixed word can combine to the right of *la-sœur*.

### 14.3.3 Prefix interpretation

According to (6), words and phrases bearing positive EDGE specifications are subject to a strict word order constraint (they can only appear as the left-most daughter in a syntactic combination). Moreover, a well-formed, complete utterance can have only negative (*none*) EDGE values, so at some point in the syntactic derivation, all positive values introduced by prefixation in the lexicon must be eliminated.

<sup>6</sup>This feature is also adopted from LKB implementation conventions. The DOM list of Reape (1994) could also be used, but the full apparatus of Domain Theory is not needed for our purposes.

The unary syntactic rules that apply in Figure 14.3 for discharging LE and  $\dot{\lambda}$  can be formulated as rewrite rules (dominating node on the left):

(7) a. LE discharge rule

$$\boxed{1}\text{NP} \left[ \begin{array}{l} \text{CONT } \boxed{3} \cup \text{def} \\ \text{EDGE} \mid \text{LE } \text{none} \end{array} \right] \longrightarrow \boxed{2}\text{N}' \left[ \begin{array}{l} \text{CONT } \boxed{3} \\ \text{EDGE} \mid \text{LE} \left[ \begin{array}{l} \text{IN } \boxed{2} \\ \text{OUT } \boxed{1} \end{array} \right] \end{array} \right]$$

b.  $\dot{\lambda}$  discharge rule

$$\boxed{1}\text{PP} \left[ \begin{array}{l} \text{CONT } \left[ \begin{array}{l} \dot{\lambda}\text{-rel} \\ \text{ARG } \boxed{3} \end{array} \right] \\ \text{EDGE} \mid \dot{\lambda} \text{ none} \end{array} \right] \longrightarrow \boxed{2}\text{NP} \left[ \begin{array}{l} \text{CONT } \boxed{3} \\ \text{EDGE} \mid \dot{\lambda} \left[ \begin{array}{l} \text{IN } \boxed{2} \\ \text{OUT } \boxed{1} \end{array} \right] \end{array} \right]$$

In fact, these rules are rather overspecified with information pertaining to the particular examples in Figure 14.3. The basic unary rules do not refer to any grammatical features; they merely execute the instructions provided by the prefix, encoded in the daughter's IN and OUT values:

(8) Unary rules for prefix interpretation

for  $P \in \{\dot{\lambda}, \text{DE}, \text{LE}\}$ ,

$$\boxed{1} \left[ \text{EDGE} \mid P \text{ none} \right] \longrightarrow \boxed{2} \left[ \text{EDGE} \mid P \left[ \begin{array}{l} \text{IN } \boxed{2} \\ \text{OUT } \boxed{1} \end{array} \right] \right]$$

The right hand side of this schema contains a cyclic structure, which is somewhat unusual but completely unproblematic from a formal point of view. If a phrase unifies with its own IN value, then the grammatical contribution of the corresponding prefix can be activated, and the phrase that was stored in OUT is realized syntactically. The resulting phrase (left hand side of the schema) always has a *none* value for the relevant prefix feature. We can now formalize the idea suggested above in §14.3.1, identifying the OUT and IN values by default. This is achieved without additional stipulation if we adopt the Generalized Head Feature Principle of Ginzburg and Sag (2001); in (8), the mother  $\boxed{1}$  will share all the features of the daughter  $\boxed{2}$  unless explicitly specified otherwise.

## 14.4 Varieties of *de*

This final section applies the phrasal affix analysis to some particular constructions involving the highly versatile element *de*. When used as a preposition (e.g., with directional semantics) or as a marker (e.g., before an AP or VP[*inf*]), *de* can be

straightforwardly handled by the approach outlined in the previous section. In other cases, a special treatment is called for.

First, the form *de* in combination with *le* can form a complex indefinite or partitive article appropriate for plural nouns or singular mass nouns (e.g., *des-petits animaux*, *de-la-bière belge*).<sup>7</sup> Morphologically, this complex article should be treated as the cooccurrence of the two prefixes *de* and *le*, but the combination is syntactically and semantically non-compositional. To account for this, we can introduce the following lexical prefixation rule:

(9) *de*-prefixation (indefinite/partitive article)

$$\left[ \begin{array}{c} \text{word} \\ \text{PHON } \boxed{1} \\ \text{EDGE } \left[ \begin{array}{c} \text{LE } | \text{ IN } \boxed{2} \\ \text{À } \text{none} \\ \text{DE } \text{none} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{c} \text{word} \\ \text{PHON } de + \boxed{1} \\ \text{EDGE } \left[ \begin{array}{c} \text{LE } \text{none} \\ \text{DE } \left[ \begin{array}{c} \text{IN } \boxed{2} N' [\text{plur} \vee \text{mass}] \\ \text{OUT NP } [\text{indef}, \text{MARKING } de] \end{array} \right] \end{array} \right] \end{array} \right]$$

This rule requires some explanation. The input is a word with a positive LE value (the output of rule (4), for example). The rule adds a second prefix *de*-, with possible idiosyncratic contraction into *du*- or *des*-, along the lines of (5). Crucially, however, this rule eliminates the positive LE specification, thus preventing any independent interpretation of the *le*- prefix in the syntax. Instead, the output has a single positive EDGE specification, for the feature DE. The IN value is identical to the original LE | IN value (which preserves the number and gender agreement between the complex article and the  $N'$ ), and the OUT value is an indefinite/partitive NP.<sup>8</sup>

A related phenomenon is the “haplology” of *de* illustrated in example (3a). Descriptively, the complex determiner *de+le* is omitted after *de* (preposition or marker). The fact that *de*- cannot occur twice (\**de-des-livres*) is easily accounted for, because all *de*-prefixation rules require an input with the feature [DE: none]. And the following lexical rule allows a single occurrence *de* to perform all of the relevant functions at the same time:

<sup>7</sup>Cases where the sequence *de+le* actually represents the preposition/marker *de* followed by the definite article *le* simply involve the successive application of two unary interpretation rules.

<sup>8</sup>The [MARKING: *de*] feature accounts for the fact that NPs containing the complex article *de+le* can be pronominalized as *en* (Abeillé et al., ms).

(10) *de*-prefixation (“haplology”)

$$\left[ \begin{array}{c} \text{word} \\ \text{PHON } \boxed{1} \\ \text{EDGE } \left[ \begin{array}{c} \text{LE } \textit{none} \\ \text{À } \textit{none} \\ \text{DE } \textit{none} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{c} \text{word} \\ \text{PHON } \textit{de} + \boxed{1} \\ \text{EDGE} \mid \text{DE } \left[ \begin{array}{c} \text{IN } \text{N}' [\textit{plur} \vee \textit{mass}] \\ \text{OUT } \text{PP} [\textit{MARKING de}] \end{array} \right] \end{array} \right]$$

The output of the rule has a single positive EDGE feature, which will trigger a unary rule application in the syntax transforming a plural or mass N' into a PP. (The CONTENT value of the PP in OUT must also specify that the embedded nominal argument is interpreted as indefinite or partitive.)

The complex article *de+le* exhibits one last alternation:

- (11) a. trouver *de-vieux* livres / *des-vieux* livres  
 b. cuisiner d'excellents plats / *des-excellents* plats  
 c. faire *de-très* bonne musique / *de-la-très* bonne musique

If the head N is not the first element in its N', the full complex article *de+le* can optionally be reduced to *de*, in particular before a vowel (in which case the reduced form is *d'*). If one assumes the traditional analysis of *de* as an independent lexical item, the constraints on this reduction are very difficult to state. The opposite is true of the phrasal affix approach, because prefixation applies precisely to the first element of N', and it is perfectly natural that the phonological and other grammatical properties of this host should influence the realization of the prefix.

The facts in (11) can be accommodated by introducing a variant of rule (9):

(12) *de*-prefixation (reduced indefinite/partitive article)

$$\left[ \begin{array}{c} \text{word} \\ \text{PHON } \boxed{1} \\ \text{HEAD } \neg \textit{noun} \\ \text{EDGE } \left[ \begin{array}{c} \text{LE } \textit{none} \\ \text{À } \textit{none} \\ \text{DE } \textit{none} \end{array} \right] \end{array} \right] \mapsto \left[ \begin{array}{c} \text{word} \\ \text{PHON } \textit{de} + \boxed{1} \\ \text{EDGE} \mid \text{DE } \left[ \begin{array}{c} \text{IN } \text{N}' [\textit{plur} \vee \textit{mass}] \\ \text{OUT } \text{NP} [\textit{indef}, \textit{MARKING de}] \end{array} \right] \end{array} \right]$$

The right hand side of this rule is identical to the right hand side of rule (9), but the input specifications are different. If the input word is not a noun—e.g., *vieux*, *excellents*, and *très* in (11)—we can obtain the effect of the complex determiner *de+le*, without actually applying *le*-prefixation beforehand ([LE: *none*] in the left hand side of rule (12)). Note that non-nouns can also undergo *le*-prefixation and then be input to the original rule (9), giving rise to the unreduced versions of (11).

As a final remark, the notion of phrasal affix is applicable to various phenomena in many languages besides French: possessive *-s* and the *a/an* alternation in

English, consonant mutation in Celtic, the definite article in Bulgarian, and case marking in Japanese, Korean, and Turkish. The results presented here demonstrate the usefulness of HPSG for modelling interface phenomena that present a real challenge for other grammar formalisms.

## Bibliography

- Abeillé, A., D. Godard, O. Bonami, and J. Tseng (ms). The syntax of French *à* and *de*: An HPSG analysis.
- Copestake, A. (2002). *Implementing Typed Feature Structure Grammars*. CSLI Publications, Stanford, CA.
- Ginzburg, J. and I. A. Sag (2001). *Interrogative Investigations: The Form, Meaning and Use of English Interrogatives*. CSLI Publications, Stanford, CA.
- Klavans, J. (1985). The independence of syntax and phonology in cliticization. *Language*, **61**:95–120.
- Miller, P. and P. Monachesi (in press). Les pronoms clitiques dans les langues romanes. In D. Godard, ed., *Les langues romanes. Problèmes de la phrase simple*. CNRS Editions, Paris.
- Miller, P. H. (1992). *Clitics and Constituents in Phrase Structure Grammar*. Garland, New York.
- Miller, P. H. and I. A. Sag (1997). French clitic movement without clitics or movement. *Natural Language and Linguistic Theory*, **15**:573–639.
- Reape, M. (1994). Domain union and word order variation in German. In J. Nerbonne, K. Netter, and C. Pollard, eds., *German in Head-Driven Phrase Structure Grammar*, volume 46 of *CSLI Lecture Notes*, pp. 151–197. CSLI Publications, Stanford, CA.
- Tseng, J. (2003a). Edge features and French liaison. In F. van Eynde, L. Hellan, and D. Beermann, eds., *Proceedings of the 9th International HPSG Conference*. CSLI Publications, Stanford, CA.
- Tseng, J. (2003b). Un traitement lexical des affixes syntagmatiques du français. In B. Fradin, G. Dal, F. Kerleroux, N. Hathout, M. Plénat, and M. Roché, eds., *Les Unités morphologiques / Morphological Units*. Silex, Villeneuve d'Ascq.
- Tseng, J. (to appear). LKB grammar development: French and beyond. In F. Fouvry, M. Siegel, D. Flickinger, and E. Bender, eds., *Proceedings of the ESSLLI-2003 Workshop "Ideas and Strategies for Multilingual Grammar Development"*. Vienna.