

- Chierchia, Gennaro (1995). *Anaphora, Presupposition and the Theory of Grammar*. Chicago: The Chicago University Press.
- Chomsky, Noam (1993). A minimalist program for linguistic theory. In *The View from Building 20*, Kenneth Hale and Samuel Jay Keyser (eds.), 1–52. Cambridge, MA: MIT Press.
- Drijkoningen, Frank (1994). Morphological strength: Basic and derived positions in French. Unpublished manuscript, Utrecht University.
- Figueiredo-Silva, Cristina (1994). La position sujet en Portugais Brésilien. Unpublished Ph.D. dissertation, University of Geneva.
- Friedemann, Marc-Ariel (1995). Sujets syntaxiques: Position, inversion et pro. Unpublished Ph.D. dissertation, University of Geneva.
- Giorgi, Alessandra and Fabio Pianesi (1994). Features scattering. Paper presented at the Conference Going Romance, Utrecht.
- Hulk, Aafke (1993). Residual verb second and the licensing of functional features. *Probus* 5: 127–154.
- Klima, Edward (1964). Negation in English. In *The Structure of Language*, Jerry A. Fodor and Jerold J. Katz (eds.). Englewood Cliffs, NJ: Prentice-Hall.
- May, Robert (1985). *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.
- Moro, Andrea (1991). The raising of predicates: Copula, expletives and existence. *MIT Working Papers in Linguistics* 15. Massachusetts Institute of Technology, Cambridge.
- (1993). *I predicati nominali e la struttura della frase*. Padova: Unipress.
- (to appear). *The Raising of Predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
- Pesetsky, David (1983). Paths and categories. Unpublished Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge.
- Pollock, Jean-Yves (1994). Notes on checking theory, pro-drop, free variation and economy. Paper presented at the Going Romance Conference, Utrecht.
- Rizzi, Luigi (1987). Three issues in Romance dialectology. Paper presented at the post-GLOW workshop on Dialectology, Venice.
- (1990). Residual verb second and the *wh*-criterion. *Technical Reports in Formal and Computational Linguistics* 1. University of Geneva.
- (1995). A note on *do*-support. Unpublished manuscript, University of Geneva.
- Sportiche, Dominique (1988). A theory of floating quantifiers and its corollary for constituent structure. *Linguistic Inquiry* 19: 425–449.
- (1992). On clitic constructions. Unpublished manuscript, University of California at Los Angeles.
- Solà, Jaume (1992). Agreement and subjects. Unpublished Ph.D. dissertation, University of Barcelona.
- Suñer, Margherita (1994). V-movement and the licensing of agreement *wh*-phrases in Spanish. *Natural Language and Linguistic Theory* 12: 335–372.
- Zubizarreta, Maria Luisa (1994). Word order, prosody and focus. Unpublished manuscript, University of Southern California, Los Angeles.

## External allomorphy and contractions in Romance<sup>1</sup>

JOAN MASCARÓ

### Abstract

*Lexical items that show different, phonologically unpredictable allomorphs whose distribution is governed by phrasal phonological conditions (like beau – bel in French) are examined. The problems that originate from their dual phonological/phrasal status are solved through the minimal assumption that if the competition between two lexical allomorphs is not resolved in the lexicon, but is decided by Evaluation at the phrasal level. Under such circumstances an effect of emergency of the unmarked arises, and as predicted, the syllabically less marked candidate is chosen. Cases of single phonological underlying representations interpreting two different lexical items (contractions) are also analyzed, and a solution which is based on the relation of phonological and other grammatical properties in the lexicon is proposed.*

### 1. Allomorphy and non-allomorphy

A pattern of emergence of unmarked phonological properties arising under special conditions (“emergency of the unmarked”) is derived by McCarthy and Prince (1994) from properties inherent to Optimality Theory. They argue that, in cases of reduplication, some normally active constraints remain exceptionally unviolated, and cause thereby the activation of lower ranked constraints that are normally inactive. In this paper I argue that a parallel pattern obtains for external allomorphs, which consist of multiple underlying forms. In the process of selecting the optimal form, this multiplicity has the effect of transferring the

1. This work was made possible by grant DGICYT PB93-0893-CO4. I am indebted to Eulàlia Bonet, François Dell, and to attendants of the Going Romance conference for their comments.

burden of choice to low-ranked, usually undecisive constraints like ONSET and NO-CODA. In the first sections I analyze cases of allomorphic variation that is determined outside the domain of the word, a topic which has been paid only sparing attention in the literature. I will then extend the analysis of external allomorphy to contractions, a related empirical domain.

In order to avoid terminological misunderstanding, I will first clarify the intended meaning of some of the terms used. *Allomorphy* will be reserved to refer to *lexical allomorphy*, that is to phonetic alternations of the same lexical unit that are not reducible to a single underlying form (as in English *sing-sang*, or *go-went* — the term *suppletion* is also used for such cases); I therefore exclude from its domain phonetic alternations of a single underlying form that can be derived by phonological rule. In normal cases of allomorphy (*plain allomorphy*) the choice of the allomorphic variant of a given morpheme depends on internal properties of the word that contains it, usually properties of other morphemes appearing in the same word. In *external allomorphy* (also referred to as *phrasal allomorphy*) the ultimate phonetic form of a word is, as in the case of *plain allomorphy*, undervivable by the sole effect of the phonology; but, in the former, the choice of the allomorph is determined outside the domain of the word, normally by properties of an adjacent word. In *contractions* a single allomorph appears which is determined lexically by, and corresponds to two different lexical categories.<sup>2</sup> In (1) some illustrative examples of each type are shown:

- |     |    |  |  |
|-----|----|--|--|
| (1) | a. | Plain allomorphy (French)                              | <i>angle – angulaire, la+Ø, le+z</i>                       |
|     | b. | Phonology (Spanish)                                    | la /__C      la casa<br>l /__a (unstressed) lØ amiga       |
|     | c. | External allomorphy,<br>phonological control (Italian) | il /__C      il cane<br>l /__V      l'amico                |
|     | d. | External allomorphy,<br>contraction (French)           | o      au temps<br>a ... l(ə) <sup>3</sup> à tout le temps |

In (1a) two roots appear, *angl-*, *angul-* whose phonetic form cannot be derived by phonological processes; it becomes necessary to list them (together with the morphological contexts to which they are associated) in the lexicon; the same applies to the feminine singular and feminine plural root of the definite article. But in the Spanish example (2b), the alternance [la]/[l] can be obtained from a regular phonological rule deleting [a] in certain contexts, for example before another unstressed [a]. Consider now (1c). It is similar to (1a) in the sense that there are no regular phonological processes of i-insertion, or i-deletion

2. Sometimes the term *portmanteau* is used in a similar sense. It was used already by Hockett (1947) for cases like the French *à + le = au* discussed below.  
3. Following common practice, I use [ə] for the partly rounded French schwa ([ɘ]); it appears also transcribed as [e] in the literature.

that could govern the alternance, deriving [il] from [l], or [l] from [il]. The alternance is hence lexical; but, contrary to (1a), the domain in which the alternance is defined is word external: the shape of the article depends on phonological properties of the following word. In (1d) the domain that governs the appearance of the allomorph [o] is also more than one word, but here a single morph appears as the phonological representation of two different words ([a] and [lə]).

## 2. Phonologically controlled external allomorphy

In many cases of external allomorphy the choice of the allomorph is determined by the phonological shape of a following word. Consider (2a), where the external allomorphs of the French prenominal adjectival form *beau* 'beautiful' (MASC SING) is shown. The form /bel/ appears whenever a vowel follows, /bo/ is found elsewhere, that is when the following segment is a consonant (including *h muet*), or in final position. In (2b) I list the other cases of prenominal elements showing the same distributional properties:

- |     |                           |                     |                    |        |               |
|-----|---------------------------|---------------------|--------------------|--------|---------------|
| (2) | French "Belle allomorphy" |                     |                    |        |               |
| a.  | /__V                      | /__C                | /__ ]              |        |               |
|     | <i>bel ami</i>            | <i>beau mari</i>    | <i>il est beau</i> |        |               |
|     | [bel ami]                 | [bo mari]           | [bo]               |        |               |
|     | 'beautiful friend'        | 'beautiful husband' | 'he is beautiful'  |        |               |
|     |                           |                     | <i>beau à voir</i> |        |               |
|     |                           |                     | [bo/#avwar]        |        |               |
|     |                           |                     | 'beautiful to see' |        |               |
| b.  | /__V                      | /__C,               | /__ ]              |        |               |
|     | <i>nouvel</i>             | [nuvel]             | <i>nouveau</i>     | [nuvo] | 'new-MASC'    |
|     | <i>vieil</i>              | [vjej]              | <i>vieux</i>       | [vjø]  | 'old-MASC'    |
|     | <i>fol</i>                | [fɔl]               | <i>fou</i>         | [fu]   | 'mad-MASC'    |
|     | <i>cet</i>                | [sɛt]               | <i>ce</i>          | [sə]   | 'this-MASC'   |
|     | <i>mon</i>                | [mɔn]               | <i>ma</i>          | [ma]   | 'my-FEM'      |
|     | <i>ton</i>                | [tɔn]               | <i>ta</i>          | [ta]   | 'your-FEM'    |
|     | <i>son</i>                | [sɔn]               | <i>sa</i>          | [sa]   | 'his/her-FEM' |

Nonallomorphic cases:

- |    |                    |                     |                 |                  |
|----|--------------------|---------------------|-----------------|------------------|
| c. | <i>joli ami</i>    | <i>joli mari</i>    | 'nice friend'   | 'nice husband'   |
|    | <i>quel ami</i>    | <i>quel mari</i>    | 'which friend'  | 'which husband'  |
| d. | <i>deu[z] amis</i> | <i>deu[Ø] maris</i> | 'two friends'   | 'two husbands'   |
|    | <i>peti[t] ami</i> | <i>peti[Ø] mari</i> | 'little friend' | 'little husband' |

Domain

- |    |   |
|----|---|
| e. | ((Mon ( <sub>D</sub> bel ami)) (est ( <sub>D'</sub> (beau)(à voir)))) |
|    | 'my beautiful friend is beautiful to see'                             |

The examples in (2c, d) represent "normal" cases. In (2c) the phonological shape of the V-final prenominal element does not vary depending on the following context. (2d) are normal cases of C-final liaison, in which the only change in the first element is the presence or absence of a final, predictable liaison consonant.

I will not address here the question of the domain in which this kind of allomorphy takes place, that is why the vowel conditioned allomorph appears in the domain **D**, but not in **D'** in (2e). Although this is an important topic, it can be kept separate from the segmental properties of the allomorphic alternations and their contexts, the issue that will be analyzed here.

The first important property to be noticed is that allomorphic cases of French are restricted to a set of lexical items, namely *beau, nouveau, vieux, fou, ce, ma, ta, and sa*. Other words similar in grammatical properties and phonological shape, like *joli, peu, quel, pareil, vrai, etc.*, do not follow the same pattern (cf. 2c), and show a single underlying phonological form.

Belle allomorphy is also morphologically restricted to specific elements in the inflectional paradigm, since only the masculine singular form is affected by allomorphic variation. The feminine and plural forms *belle, belles, beaux* show phonetic variants that can be solely derived from phonological processes, such as *liaison*. Thus the minimal lexical specifications for *beau* (3a) must state that it has two allomorphic phonological variants, and that one of them, /bo/, corresponds to the masculine, and that the other, /bɛl/, corresponds to both the masculine singular and to the feminine. From this representation we can derive morphologically the inflected forms of the adjective, shown in (3b), that is the inflective paradigm of *beau* (zero morphs are not indicated):

(3) a. Lexical representation of *beau*

/bo/ - [-FEM]  
/bɛl/ - [+FEM], [-FEM, -PL]

b. Inflectional paradigm

	Masculine	Feminine
Singular	/bɛl/, /bo/	/bɛl+V/
Plural	/boz/	/bɛl+V+z/

In trying to account for the distribution of [bo] and [bɛl], and other cases of phonologically controlled external allomorphy, we face a paradox. Since the phonological context that controls the allomorphic choice is completely regular (it is, for instance, the same as for *liaison*), we might be tempted to try to derive the allomorphic distribution of [bɛl] and [bo] and similar cases phonologically. But the phonetic contrasts that obtain in this regular context ([ɛl]~[o], [ø]~[ɛj], [u]~[ɔl], etc.) show that the *phonetic alternations are totally idiosyncratic* (unlike *liaison*), and, as already mentioned, *lexically reduced*, that is limited to a small set of lexical items; they are not derivable through phonological processes.

Should we insist on a phonological treatment, then we would be obliged to mark these few lexical items with a lexical diacritic, which would trigger a totally implausible phonological process. This seems to indicate that the solution must be lexical, more specifically allomorphic: the lexicon should contain more than one single underlying form. If we choose an allomorphic solution, however, we are led to two undesirable consequences. First, in "normal" cases allomorphy is determined within the word domain, it takes place prior to lexical insertion. But the expressions (*bɛl ami*), (*nouvel âge*), etc. are not lexical items, hence the lexicon would have to be extended to include these domains. Second, if we adopt a lexical solution we are removing the expression of a phonological regularity from the phonology, since the context that determines the allomorphic choice is phonologically regular.

The case of Belle allomorphy is not isolated. In most Romance languages we can find similar examples. Although each case needs an individual, careful analysis, an overview of similar instances within Romance is illustrative. In (4) a selection of parallel cases is presented. In all these cases, like in French Belle allomorphy, the phonological regularity that governs the choice of the allomorph involves the sapa of the allomorph and the shape of the adjacent word at the margins at which both make contact, as in *bɛ[l a]mi* and *b[o m]ari*. Examples in (4) are grouped according to the phonological shape of the allomorph at this margin:<sup>4</sup>

(4) Some representative cases of phonologically conditioned external allomorphy:

Context Shape	a.		b.	
	___C	___V	___C	___V
...CV	...CV	...C	...VC	...C
<i>la</i>	<i>l</i>		<i>il</i>	<i>l</i>
<i>de</i>	<i>d</i>		<i>en</i>	<i>l</i>
<i>di</i>	<i>d</i>			
<i>se</i>	<i>s</i>			
<i>que</i>	<i>qu</i>			
<i>mi</i>	<i>m</i>			
<i>ti</i>	<i>t</i>			
<i>si</i>	<i>s</i>			
<i>ne</i>	<i>n</i>			

4. Other cases are not difficult to find. The examples in (4) correspond to Italian (article: *la, il*, preposition: *di, ad*, pronominal clitic: *mi, ti, si, ne*), Occitan (conjunction: *que, se*, preposition: *de, a/an*), Catalan (article: *en/l*, pronominal clitic: *ho*), French (article: *la*, adjective: *bɛl*). For the Calabrese *chid* 'what' (and Neapolitano and Romanesco *ched*), see Rohlfs (1966: § 489).

c.		d.	
$\_V$	$\_C$	$C\_C$	$V\_V$
...VC	...V	...V...	...C...
<i>bel</i>	<i>beau</i>	<i>ho</i> [o]	<i>ho</i> [w]
<i>ad</i>	<i>a</i>		
<i>chid</i>	<i>chi</i>		
<i>an</i>	<i>a</i>		

In all the cases in (4) the context exhibits a phonological regularity that is similar to the one we have seen in French Belle allomorphy. Such cases present thus a theoretical problem, because they do not lend themselves to a phonological or to a morphological solution, and yet they are not singularities that might be attributed to error of analysis or to high lexical idiosyncrasy, since they appear to be quite widespread, and they present always similar characteristics.

Let us give a closer look to the phonological conditions. One crucial aspect is that in phonologically controlled external allomorphy the choice of an allomorph is not based on the fact that the other allomorph is rejected for reasons of absolute phonological ill-formedness. For instance, \*[bel mari], \*[bo ami] are ungrammatical, but not because they are phonologically impossible (cf. phonologically similar forms like *quel mari* [kel mari], *joli ami* [zoli ami]). Yet, they are not parallel to other cases of plain allomorphy, like *angle - angulaire*, for which the contexts determining the choice are phonologically random. Nor are they reducible, as we have already seen, to regular phonological cases for which the context draws the line between what is possible and impossible phonologically, as in liaison: *deuØ maris*, *deu[z] animaux*, \**deu[z] maris*, \**deuØ animaux*.

More interestingly, the examples in (4) can be shown to obey the same phonological conditions. If we were to propose a descriptive approximation to this phonological regularity, we might say that the allomorphic options chosen are, in some sense, syllabically less marked. Therefore [bo mari] wins over [bel mari], because its CVCV structure is less marked than the structure CVCCV of the latter. Similarly, [bel ami] wins over [bo ami], since CVCV... is less marked than CVV... Let us state this generalization somewhat more precisely:

- (5) The allomorph chosen is the one which can be best syllabified with adjacent words to yield an output closer to a core syllable (that is CV).

The validity of (5) can be easily checked against the data in (4).

It should be obvious that there is no way to filter out the incorrect allomorphic choice by resorting to the normal phonology of the language. We would be forced to appeal to some stipulation relating general properties of markedness to allomorphic choice.

Interestingly enough, Optimality Theory (Prince and Smolensky 1993) makes very specific predictions in this case. Let us see why and how.

Unmarked structure is expressed by universal constraints. This does not mean that marked structures are absolutely prohibited, which is evidently false. Marked structures can be generated by a given grammar G, and they are marked because constraints expressing a certain type of unmarkedness are dominated by higher ranked constraints that license the marked structures. Thus satisfaction of higher ranked constraints licenses structures even if they violate lower ranked constraints. But, crucially, these lower ranked constraints are *still* part of the grammar G. In the case of syllabic structure, a core syllable CV is unmarked because there is a constraint, ONSET, disallowing empty onsets and a constraint prohibiting full codas, NO-CODA. If a language allows for more marked syllabic structures, like VC, this will be because the effect of ONSET and NO-CODA will have been superseded by other, higher ranked constraints. ONSET and NO-CODA, and other constraints to be referred to later, are defined below:

- (6) ONSET: Syllables must have onsets.  
 NO-CODA: Syllables must not have a coda.  
 FILL: Epenthetic structure (□) is prohibited.  
 \*COMPLEX: No more than one C, V can associate to one syllable node.  
 PARSE-SEG: Unsyllabified segments (< >) are prohibited. [N.B., segments that remain unsyllabified are finally deleted]  
 \*COMPL-ONS: (Complex onsets must have increasing sonority.)

Underlying expressions are subject to Gen, a function that supplies each underlying expression with possible analyses (in our case, possible  $\sigma$ -structure analyses, free epenthesis and deletion), yielding as a result a set of alternative candidates. Thus, given an input *a* (an underlying expression),  $\text{Gen}(a) = \{a_1, a_2, \dots, a_n\}$ , where each  $a_i$  is a possible analysis of *a*. These multiple outputs of Gen are then subject to Evaluation (Eval). Eval chooses among the candidates in  $\{a_1, a_2, \dots, a_n\}$  some  $a_k$ , the optimal candidate, according to satisfaction and violation of constraints and their hierarchical order:  $\text{Eval}\{a_1, a_2, \dots, a_n\} = a_k$ .

Normally, the multiplicity of forms subject to Evaluation,  $\{a_1, a_2, \dots, a_n\}$ , among which Eval selects ultimately the grammatical form, will be solely due to Gen. Different grammars will have different rankings of universal constraints. Consider just an illustrative example. In Catalan the consonant contact across word boundary, C##C, in particular for C=[-cont], can have three different results depending on the dialect. Alguerese inserts a vowel, Valencian deletes the first C, and other varieties do not change syllabic structure. Thus, as shown in (7a), the form *fet bé* 'done well' ends up, respectively, as [fet e be], [fe be], [feb be], in this last case after place assimilation. For all these varieties the underlying form is the same, /fet be/.<sup>5</sup> From these identical underlying structures

5. In isolation these forms appear as [fet] and [be] in all varieties. I ignore here the (irrelevant) possibility of [be] being underlyingly /ben/.

Gen generates the same set of alternative forms. But a different ranking of constraints results in the choice of a different optimal candidate. (7b) illustrates the case in which the output remains faithful to the underlying form, even though it violates the lower ranked NO-CODA. Notice that what is crucial is that NO-CODA be ordered last; the ordering between PARSE-SEG and FILL does not matter:

(7)	a.	Underlying /fet be/	Alguerese fet e be	Valencian fe be	Other varieties fet be (→ feb be)
	b.		PARSE-SEG	FILL	NO-CODA
	i.	fet .be			*
	ii.	fe .t □ .be		*!	
	iii.	fe <t> .be	*!		

If NO-CODA precedes either PARSE-SEG or FILL, the final form will have to be unfaithful to the underlying form in order to avoid the *t* in coda position. Under such a situation, if FILL is ordered last, (7b-ii) will be chosen, and if PARSE-SEG is ordered last, the selected form will be (7b-iii).

### 3. External allomorphic choice as emergence of the unmarked

Normally, the multiplicity of forms in the set  $\{a_1, a_2, \dots, a_n\}$  subject to Evaluation is a function of an input, an underlying lexical expression *a*, and Gen. But in special cases this multiplicity of forms will emanate from different sources. This happens typically when a single lexical form *a* is not the only source of input to Gen. One such case, dealt with in detail by McCarthy and Prince (1994), has to do with reduplication. In reduplication, a specific morpheme does not pick up its content directly from the lexicon, but gets it indirectly, via an operation of correspondence that gives rise to different candidates extracted from a base. In such cases the multiplicity of forms is due to Gen and to the morphological operation.

Since there is a family of constraints, faithfulness constraints, that relates candidates to their underlying form, those candidates whose origin is not the lexicon, but which arise through a morphological operation will trivially obey faithfulness. This determines that lower ranked constraints can become crucial in determining the optimal choice, and hence an effect of emergence of unmarked structures obtains.

A simpler case of cancellation of faithfulness requirements can be traced back to the fact that a lexical item may contain more than one underlying form. External allomorphy, the phenomenon under discussion, is an instance of

multiplicity of lexical phonological forms. Normally, a lexical item *L* has a single phonological underlying form /*a*/ which constitutes the input to Gen. Gen determines the set of candidates  $\{a_1, a_2, \dots, a_n\}$ . Assume now a lexical item *L* whose phonological structure  $\Phi$  consists of two alternate underlying forms /*a*, *b*/. The optimal form is defined as the result of a single evaluation of all the outputs contained in *Gen*(*a*) and in *Gen*(*b*), that is, Eval applies to the union of the two outputs of Gen (this can of course be extended to more than two underlying forms):

- (8) For a lexical item *L* such that  $\Phi = a, b$ :  
Eval (Gen (*a*, *b*)) = Eval (Gen (*a*)  $\cup$  Gen (*b*))

If allomorphic choice between *a* and *b* is determined morphologically (the normal case in *internal allomorphy*), the selection will be done in the lexicon before phonological evaluation. But if not, then given (8), Optimality Theory dictates that the choice between *a* and *b*, that is the choice between some Gen (*a*) and some Gen (*b*), is determined by the function Eval:

- (9) (External) allomorphic choice is determined by Evaluation.

To see the consequences of evaluating the sum of two (or more) underlying forms, consider a hypothetical example, based on the data in (7a). In (7b) above, the ranking of constraints PARSE-SEG, FILL >> NO-CODA, determines faithfulness to the underlying representation /fet/. Suppose we add now ONSET to the constraints, and that the ranking is PARSE-SEG, FILL >> NO-CODA, ONSET. Consider two hypothetical underlying representations, /fet/ and /fe/. Since the faithfulness constraints are ordered before the constraints relating to  $\sigma$ -structure, the optimal candidates will be the faithful outputs of Gen [fet] and [fe], respectively.

But assume now that /fet/ and /fe/ are part of the same underlying representation, that is they are allomorphs. Gen (/fet/) and Gen (/fe/) will now both compete for optimal evaluation. Since the outputs of Gen ([fet]) and ([fe]) both comply with faithfulness requirements ([fet] is faithful to /fet/ and [fe] is faithful to /fe/), they will both satisfy PARSE-SEG and FILL: Faithfulness cannot determine the choice. This means that the lower ranked constraints, NO-CODA and ONSET, now become crucial in deciding the optimal candidate. Both candidates satisfy ONSET, since both have an onset, but /fet/ has a coda, thus violating NO-CODA. Thus in principle [fe] should be chosen over [fet]. At the word level, this would mean that [fet] could never be chosen, hence it would be an impossible underlying representation for *fet*.<sup>6</sup> But if the domain of

6. This relates to the problem of the prosodic domain of external allomorphy. It might explain why external allomorphy appears most commonly in lexical items that do not show up in isolation, such as prepositions, determiners, and pronominal adjectives.

Evaluation is higher, that is if we consider a constituent that contains both *fet* and the following word, NO-CODA and ONSET will have different effects depending on the context. If a vowel follows, the final *t* of [fet] will be able to syllabify with it, whereas the final vowel of the other candidate, [fe], will cause hiatus. If a consonant follows, the final *t* will be syllabified as a coda. The tableaux (10) and (11) show the evaluation of our hypothetical example in these two phrase level contexts:

(10) Hypothetical example. Multiple underlying form: /fet, fe/

Evaluation /\_\_V

	PARSE-SEG	FILL	NO-CODA	ONSET
a.	☞ fe .t V...			
b.	fe .V...			*!

(11) Evaluation /\_\_C

	PARSE-SEG	FILL	NO-CODA	ONSET
a.	fe .C...		*!	
b.	☞ fe .C...			

In both (10) and (11) the higher ranked constraints, PARSE-SEG, and FILL, are satisfied, and they are unable to determine the choice of the optimal candidate. This causes lower ranked constraints, that normally do not have a chance to play a role in evaluation, to *emerge*, dictating preference for the unmarked structure. Thus evaluation is governed here by NO-CODA and ONSET, as shown in (10) and (11). Before a vowel the allomorph [fet] is chosen because CV.CV... is preferred over CV.V... which lacks an onset. Before a consonant, the allomorph [fe] is chosen: here CV.C... is preferred over CVC.C..., since the latter contains a syllable final C that violates NO-CODA.

Thus the existence of *two* lexical sources subject to the *same* evaluation function determines an emergence of the unmarked effect. Returning now to our real cases of external allomorphy in (4), the unmarked effect that was summarized in (5) can be obtained in the same way as for our hypothetical [fet]/[fe] example. Consider the French case. Unlike most inflected lexical items, *beau* will have two alternative phonological interpretations: {bel, bo} (Free variation will be discussed in section 6.) Since *beau* has two underlying interpretations at the phonological level, Gen will apply to both to yield two sets of candidates. Eval will apply to the union of the sets, as indicated in (8). For reasons of simplification, we can determine the optimal element in Eval (Gen

{bel, bo})) = Eval (Gen (bel) ∪ Gen (bo)), by determining Eval (Eval (Gen (bel)), Eval (Gen (bo))). In other words we determine the best candidate by picking up the optimal output of Gen for /bel/ and the optimal output of Gen for /bo/, and then we compare them to determine which one is selected as optimal by Eval. In (12a-c) we examine the choice by evaluating the optimal output of Gen for each of the two allomorphs in each of the three contexts that condition the choice. They correspond to *bel ami*, *beau mari*, and *il est beau* in (2):<sup>7</sup>

(12) a. Evaluation /\_\_V

	ONS	NO-CODA
☞ be .l a .mi		
bo .a .mi	*!	

b. Evaluation /\_\_C

	ONS	NO-CODA
bel .ma .ri		*!
☞ bo .ma .ri		

c. Evaluation /\_\_]

	ONS	NO-CODA
bel		*!
☞ bo		

#### 4. Similar cases

In this section I examine three other cases of phonologically controlled external allomorphy taken from Catalan. Consider first the alternation of the personal definite article. In several varieties of Central Catalan two definite articles are distinguished: the *common* definite article and the *personal* article. The latter is

7. How much this analysis carries over to internal allomorphy must remain an open question. For Romance languages it seems to be the case that, as noted before, internal allomorphy is highly lexical. Nevertheless it appears that at least some cases that are neither lexical entirely nor purely phonological, in the usual sense, lend themselves to a similar analysis. This is the case of the negative prefix *a-/an-*, in several Romance languages. *A-* appears before consonants and *an-* before vowels, so as to give a V.C..., and avoiding a V.V..., VC.C... syllabification. Some examples from Catalan are: *an+aerobi*, *an+ovulatori*, *an+històric*, *an+hidre*, *an+harmònic*, *an+arquia*, *an+alfabet*, vs. *a+cromàtic*, *a+normal*, *à+crata*, *a+fàsia*, *a+fònic*, *à+graf*, *a+moral*, *a+morf*, *a+rrítmic*, *a+sèpsia*, *a+sexual*, *a+sil·làbic*, *a+simètric*, *a+teisme*, *à+ton*.

restricted to unique human proper names in the masculine singular. The *personal* article takes two forms, *en* before consonant initial nouns, and *l'* before nouns beginning with a vowel. (13a) shows some examples of the personal article, and (13b) some examples of the common article:

- (13) a. *l'Emili*            *en Juli*  
           *l'Otto*             *en Fritz*  
  
           *l'Einstein*       *en Wittgenstein*  
           *l'Alan Turing*    *en Turing*  
           *l'Allende*        *en Salvador Allende*
- b. *l'home*            *el jove*  
    'the man'        'the youngster'  
  
    *l'amic*            *el primer Wittgenstein*  
    'the friend'     'the first Wittgenstein'
- c. Lexical representation of the article:  
    /l/: common definite, personal definite  
    /ən/: personal definite

The personal article has thus two forms. One is *en*, the other is identical to the corresponding form of the common definite article, *l'*. Thus whereas *en Wittgenstein* has only a unique interpretation, and *el Wittgenstein* only a non-unique one (as in *el primer Wittgenstein* 'the first Wittgenstein'), *l'Einstein* has both interpretations. The phonological lexical properties of the article are given in (13b).

Let us examine first the case of the common article. Here the two phonetic forms, /l/ and /əl/, can be regularly derived from a single underlying form /l/ by a process of epenthesis that provides an extra initial vowel to words beginning in clusters. Since FILL must be violated (there *is* epenthesis) a higher constraint must license the epenthesis structures. Simplifying the matters slightly, I will use \*COMPL-ONS as the constraint(s) whose effect is to disallow certain onset clusters, namely those with nonincreasing sonority. Peripheral epenthesis in Catalan is not confined to this particular lexical item, but is totally general whenever unpermissible CC sequences arise initially. \*COMPL-ONS will outrank FILL, ONSET, and CODA, and it will be violated by unepenthesis articles, thus allowing for the epenthesis form, which (irrelevantly) violates all three lower ranked constraints. The examples [l ɔ .mə], [əl ʒoβə] in (14) are taken from (13).

- (14) a. Evaluation /\_\_C
- |           | *COMPL-ONS | FILL | ONSET | NO-CODA |
|-----------|------------|------|-------|---------|
| l ʒo .βə  | *!         |      |       |         |
| əl ʒo .βə |            | *    | *     | *       |

- b. Evaluation /\_\_V
- |          | *COMPL-ONS | FILL | ONSET | NO-CODA |
|----------|------------|------|-------|---------|
| əl ɔ .mə |            |      |       |         |
| l ɔ .mə  |            | *!   | *     |         |

As opposed to the nonpersonal definite article, the personal article has two allomorphic variants. One surfaces as [ən] always; we therefore posit an underlying structure /ən/. The other is identical to the common article. I am assuming here (crucially, for this case) the lexical structure suggested in (13c), that is, that in those cases in which two forms in a paradigm share a phonologically identical form, there is a single morph shared by the two elements in the paradigm. Since, as we have seen above, the phonological form of the common article is /l/, the allomorph of the personal article will be /l/ also. Thus the result of applying Gen to the personal article, Gen (*personal article*), will be Gen (/l/, /ən/).

Evaluation of the allomorphs is shown in (15). Although we can just compare the optimal result of one of the allomorphs (14) to the optimal result of the other allomorph (cf. (8) above), we compare all relevant forms together in (15). Before a consonant (15a) the epenthesis [əl] (= □l) will fare better than [l], as already shown in (14) above. Unepenthesis /ən/ will better satisfy the set of constraints, because [əl] violates FILL. Before a vowel (15b) the best candidate of Gen (/l/) will be unepenthesis [l], as shown already in (14). The best candidate of Gen (/ən/) will be the faithful form [ən]. But [ən] will be chosen by Eval over [l], because it yields a CV structure, whereas /ən/ presents a VC structure that violates the constraint ONSET:

- (15) a. Evaluation /\_\_C
- |           | *COMPL-ONS | FILL | ONSET | NO-CODA |
|-----------|------------|------|-------|---------|
| ən ʒu .li |            |      | *     | *       |
| l ʒu .li  | *!         |      |       |         |
| □l ʒu .li |            | *!   | *     | *       |
- b. Evaluation /\_\_V
- |                | *COMPL-ONS | FILL | ONSET | NO-CODA |
|----------------|------------|------|-------|---------|
| ə .n ə .mi .li |            |      | !*    | *       |
| əl ə .mi .li   |            |      |       |         |
| □ .l ə .mi .li |            | *!   | *     |         |

Another variety of Catalan, Northwestern Ribagorçan furnishes one more case of external allomorphy.<sup>8</sup> The first and second person demonstratives [iʃo], [ésto] appear in this form preconsonantly and finally, but when a vowel follows the consonant final versions [iʃ], [ést] show up. No final *o* deletion before vowel occurs in other lexical items (cf. [díʃa l áltro astí] 'leave the other here'), which shows that the alternation cannot be gotten through phonological deletion.

(16)	/__C, /__]	iʃo 'that'	iʃo tʃikót 'that boy'
		ésto 'this'	ésto líβre 'this book'
	/__V	iʃ 'that'	iʃ animál 'this animal'
		ést 'this'	ést óme 'that man'

All demonstrative forms violate ONSET, and *est(o)* violates also NO-CODA. This is however a legal violation in the phonology of Catalan, since it allows for satisfaction of faithfulness to the underlying representation. But since we have two underlying phonological representations for each of the lexical items, that is {/iʃo/, /iʃ/} and {/ésto/, /ést/}, the choice of the second allomorph, shown in (17a-i) for the first example, results in an additional violation of ONSET, whereas the first allomorph avoids it through the syllabification of its final consonant with the initial vowel of the following word (17a-ii). Before a consonant, or finally (17b), it is the first form, /iʃo/, which is preferred to /iʃ/, since the latter presents an additional violation of NO-CODA:

(17) a. Evaluation /\_\_V

		ONS	NO-CODA
i	í .ʃo .óme	* *!	
ii	í .ʃ óme	*	

b. Evaluation /\_\_C, /\_\_]

		ONS	NO-CODA
i	í .ʃo .tʃ i.kót	*	*
ii	iʃ .tʃ i.kót	*	**!

The second example, {/ésto/, /ést/} is only slightly different. Before a vowel ONSET will be the decisive constraint, and the violation of NO-CODA will have no effect. Before a consonant [ést(o)], will always violate NO-CODA. Here \*COMPLEX will decide between the allomorphs, since syllabification of the last [t] of [ést] with the following vowel eliminates the complex coda:

8. Data are taken from Sistac (1993: 157-159).

(18) a. Evaluation /\_\_V

		ONS	*COMPLEX	NO-CODA
i	és .to .ó .me	* *!		*
ii	és .t ó .me	*		*

b. Evaluation /\_\_C, /\_\_]

		ONS	*COMPLEX	NO-CODA
i	és .to .tʃi .kót	*		**
ii	ést .tʃi .kót	*	*!	**

The last case to be analyzed involves both a left and a right context. In many varieties of Western Catalan (and in Old Catalan as well), the neuter clitic *ho* takes two different phonetic forms: it appears as [w] after vowels or before vowels, and as [o] elsewhere, that is between consonants, initially before C, and finally after C. In these varieties unstressed *o* does not reduce to *u*, nor is there any way to derive the [w]-[o] contrast phonologically. Notice also that here, unlike in the *beau/belle* and *en/l'* cases, none of the allomorphic forms is identical to any other form in the inflectional paradigm.

(19) contains an example of the grammatical form for each possible context; the two last columns show, respectively, the relevant syllabic structure of the grammatical form and that of the ungrammatical form which results from choosing the other allomorph:



(19) Western Catalan neuter clitic *ho*:

Context	Grammatical form	Gloss	✓ Allomorph ✓ σ-structure	*Allomorph *σ-structure
a. C__C	com [o] fa	'how does it'	ko .mo.	komw.
b. V__V	qui [w] ha fet	'who has done it'	ki .wa.	ki .o .a.
c. C__V	com [w] ha fet	'how has done it'	kom. wa.	kom. o .a.
d. V__C	qui [w] fa	'who does it'	kiw.	ki .o.
e. __V	[w] ha fet	'has done it'	wa.	o .a.
f. V__	porta [w]	'bring it!'	taw.	ta .o.
g. __C	[o] fa	'does it'	o .fá.	wfá.
h. C__	fez [o]	'do it!'	fe .zo.	fezw.
a. V__	porta [w]	'bring it!'	taw.	ta .o.
b. V__C	qui [w] fa	'who does it'	kiw.	ki .o.
c. C__V	com [w] ha fet	'how has done it'	kom. wa.	kom. o .a.
d. C__C	com [o] fa	'how does it'	ko .mo.	komw.
e. V__V	qui [w] ha fet	'who has done it'	ki .wa.	ki .o .a.
f. __V	[w] ha fet	'has done it'	wa.	o .a.
g. C__	fez [o]	'do it!'	fe .zo.	fezw.
h. __C	[o] fa	'does it'	o .fá.	wfá.

The relevant constraints are \*COMPL-ONS, \*COMPLEX, ONSET, NO-CODA. The first one, \*COMPL-ONS, dominates the rest, since it dominates FILL. ONSET dominates NO-CODA (see below). This takes care of the first three cases (19a, b, c). Some result in a violation of NO-CODA, but the other allomorphic choice violates ONSET. The next four cases (19d–g) show a maximally unmarked CV structure whereas the choice of the alternative allomorph would determine a structure containing a complex consonant cluster violating NO-CODA and \*COMPLEX in the case of (19d, g), and a structure with two empty onsets in the case of (19e, f). Finally, in (19g) ONSET is violated, but the alternative allomorphic choice violates \*COMP-ONS, a higher ranked constraint.

The ordering ONSET >> NO-CODA could be based on general grounds if it were the case that there is a general preference for CVC syllables over V syllables. Although language particular evidence for this ordering is difficult to get, precisely because its effect only "emerges" in restricted domains like the one that is occupying us now, there is however some independent evidence available. Several morphological processes of word formation studied in Cabré (1993) (see also Cabré and Kenstowicz 1995) yield structures containing typically CV(C) syllables. This suggests that constraints related to truncation outrank NO-CODA, thus allowing closed syllables, but are themselves outranked by ONSET, with the effect of banning onsetless truncated forms. In hypocoristic

truncation, for instance, full codas are permitted (20a), while onsetless V(C) are disallowed (20b); truncated hypocoristics appear in square brackets:

- (20) a. An[toni] Jose[finá] Mi[quel] Ra[mon]  
Bar[tomeu] Rai[mundo] Fran[cesca]  
b. Llu\*[ísa] Juli\*[ana] Edu\*[ard] Jo\*[an] Jo\*[aquim]

## 5. Other analyses

In this section I discuss a couple of alternative analyses of phonologically conditioned external allomorphy. One possible way to account for these cases is to codify the phonological context into the lexical representation. This is the option chosen by Piera (1985: 294–295), and by Hayes (1990). In (21) I give one example from each author. (21a) shows that Italian *i* is subcategorized as appearing before a consonant (C) or a consonant–liquid cluster (CL), *gli* appearing elsewhere. In (21b) English *an* is subcategorized as appearing before a vowel, and *a* appears elsewhere.

- (21) a.  $i$  [+ART, +MASC, +PL, ..., \_\_\_C (L) ]      Italian *i/gli* allomorphy,  
Piera (1985)
- $gli$  [+ART, +MASC, +PL, ... ]
- b.  $\left[ \begin{array}{l} 999 \\ \text{(syntactic and semantic specifications)} \\ \text{an in the context / \_\_ V} \\ \text{ə elsewhere} \end{array} \right]$       English *a/an* allomorphy,  
Hayes (1990)

This kind of approach faces a clear problem of excess of idiosyncrasy. Both the phonological shape of the allomorphs and the phonological context that governs its distribution are codified lexically. But whereas the phonological shape of the alternant allomorphs *is* idiosyncratic, the context that governs their distribution *is not*, as has been argued before. If we were to adopt such a lexical subcategorization solution, it would be equally expectable to get, for instance, *a* before vowels, and *an* before consonants. This does not amount to denying the possibility that cases of idiosyncratic contexts of distribution exist (a clear example would be of course word-internal allomorphy), but it appears that at least the vast majority of cases of phonologically conditioned external allomorphy is governed by the kind of phonological context that can be accounted for in the way proposed here.

In a different line of argumentation, Tranel (1994) presents, for the French

cases exemplified in (2), an analysis couched in Optimality Theory terms which covers at the same time liaison cases. The core of his proposal is the observation that in the French cases the inflected form that occurs before vowels is identical to the form corresponding to the opposite gender value. Thus in cases like masculine *bel ami* the allomorph chosen is the same as in feminine *belle plante*. And for feminine *mon arme* the form appearing instead of the expected *ma* is identical to the masculine *mon* as in *mon mari*. In order to capture this generalization, he proposes a constraint, GENDER, which "make[s] a determiner agree in gender with its noun." Both masculine and feminine forms are candidates to go with a following word; but violation of GENDER will disallow cases like \**bel* [+FEM] *mari* [-FEM], in favor of *beau* [-FEM] *mari* [-FEM], whereas ONSET will disallow cases like \**beau ami*. Since ONSET precedes GENDER in the ranking, *be.l* [+FEM] *a.mi* [-FEM] will be favored over \**beau* [-FEM] *.a.mi* [-FEM].

The analysis poses problems of different nature. First, phrases containing disagreeing pairs like *bel* [+FEM] *ami* [-FEM] will not be able to be licensed by the syntax — unless some extra stipulation is introduced that gets around the general conditions on syntactic agreement.

On the other hand, the allomorph-gender parallelism does not always hold: not all cases of Belle allomorphy show prevocalic forms that show the same phonological form as the allomorph of the opposite gender. For *cet/cette* (see (2)), as in *cet ami/cette femme* we would expect identical forms, but *cet* is [s(ə)t] and *cette* is [set]. Similarly, other allomorphic cases with the same structure should obey the gender generalization. Consider the case of *favori-favorite*. In French final consonants that appear in the feminine (*peti[t]e*) also appear in the masculine in liaison contexts (*peti[t] ami*). When this does not occur, as for *favori-favorite*, we will have two allomorphs, [favori] [-FEM], and [favorit] [+FEM]. We would expect the choice of the feminine form [favorit] to be extended to the prevocalic masculines. But we get *favori ami*, and not \**favori [t] ami*. Under the present analysis, the allomorphy is lexically tied to gender; the masculine has the form /favori/, the feminine the form /favorit/. Since no single inflected form will have two underlying phonological forms no external allomorphic effect is predicted.

The analysis of Belle allomorphy in terms of violation of gender agreement seems nonetheless tempting in the view that in these cases the masculine/feminine prevocalic allomorph coincides, largely, even if not generally, with the feminine/masculine normal allomorph, respectively. But this is not surprising in the French case, for two reasons. First it is generally the case that in lexical allomorphy the number of allomorphs within an inflectional paradigm is minimized, and that the same allomorph is shared by different inflectional classes (consider, for example, the distribution of the allomorphs of the verb *aller*: *vai-*, *va-*, *all-*, *i-*, *aill-*). Second, the phonological context that distinguishes

Belle allomorphy forms, namely prevocalic versus non-prevocalic, is the same as the context distinguishing feminine and masculine. Thus the identity of allomorphic forms and forms of the opposite gender is synchronically an accident, although diachronically it is clearly not a coincidence. The rest of the Romance cases that we have discussed above show that when other conditions obtain, gender (or other inflectional features) do not play a role.

## 6. Lexically controlled external (or phrasal) allomorphy: Contractions

In many cases the existence of allomorphic irregularity which is determined outside of the domain of the word is different from the examples examined so far. In some cases the choice of the variant is not subject to phonological control.<sup>9</sup> In a specific set of cases the control is purely lexical, and mutual, that is in such instances a single, irregular morph is chosen that corresponds to two specific, different lexical items. The most common case is preposition + definite article, but other cases can also be found. Several terms appear in the literature to designate this phenomenon: *contraction*, *preposizione articolata*, *Verschmelzung*, *portmanteau word*; I will use the first one. The analysis of contraction is important for external allomorphy, because many lexical items present contraction and phonologically controlled external allomorphy effects simultaneously, but I will not deal with those more complex cases here.

Compare, as an illustration, the French and Spanish reflexes of the preposition corresponding to 'to' and the definite masculine singular article, when separated by another lexical item (22a) and when in contact (22b):

(22) French and Spanish *a* contraction:  $\grave{a} + le = au$ ,  $a + el = al$

a. Separate		b. Adjacent	
<i>\grave{a} tout le</i>	<i>a todo el</i>	<i>au</i>	<i>al</i>
[a...l(ə)]	[a...el]	[o]	[al]
'to all the'		* <i>\grave{a} le</i>	* <i>a el</i>
		'to the'	

Under adjacency the regular phonological interpretation of the underlying sequences P+Article \*/al(ə)/ and \*/aɛl/ is ungrammatical. The grammatical forms, [o] and [al], respectively, cannot be obtained via phonological rules with the effect *al* → *o*, *ae* → *a* (cf. for Spanish *para el* [paraɛl], *a entrar* [aɛntrár]). The two cases are not completely parallel. For Spanish it is possible to put

9. I will leave aside the cases of possible syntactic control, like the analysis of Hausa vowel shortening before full object NPs in Hayes (1990: 93–95), which deserve more careful empirical and theoretical scrutiny.

forward an analysis where only the article varies allomorphically between /e/ and /l/. In the case of French there is no way to split the phonological form /o/ in two parts, such that one is identical to the normal form of the preposition or the article. In both cases the elements that determine the allomorphic variant are only lexical: the contraction takes place when and only when two specific lexical items, that is, the preposition 'to' and (some forms of) the definite article are adjacent. Since there is no clear argument to treat such cases differently, I will assume that both are examples of *contractions*, in the sense that there is a single phonological form, a single morph, which corresponds to two different lexical items: in (22) French /o/ corresponding to à + le, and Spanish /al/ corresponding to a + el.

Typically, in such cases the phonological form corresponding to these two lexical items takes phonologically the shape of a single word (or clitic), and adjacency is always preserved. Let us compare these cases with phonologically controlled external allomorphy. As for regular properties, contractions do not appear to be dependent on any phonological regularity; in particular, it is not the case that the contracted form systematically improves in syllabic unmarkedness with respect to the noncontracted sequence. On the other hand, they share the rest of the idiosyncratic, lexical properties of the phonologically conditioned cases: *phonetic alternations are totally idiosyncratic*, and the phenomenon is *lexically reduced* to a small set of lexical items (pairs of lexical items, in the case of contractions). Thus for contractions the inexistence of phonological regularities eliminates one of the problems faced by phonologically controlled allomorphy. On the other hand the fact that two different lexical items share a single phonological form introduces a new problem, given standard conceptions of the structure of the lexicon.

Before discussing a solution, we illustrate contraction with a few more examples:<sup>10</sup>

(23)	Language	Categories	Isolated forms	Contracted form	
	Galician	P + Art 'to'	a	o	o
			a	os	os
			a	a	ã
	French	P + Art 'of'	a	as	ãs
			d(ə)	l(ə)	dy
			d(ə)	le(z)	de(z)

10. The phenomenon is of course not restricted to Romance. Cf. the German "Verschmelzungen" like *an dem* → *am*, *zu der* → *zur*, *in dem* → *im*, etc. or the English forms *is not* → *ain't*, *want to* → *wanna*, etc.

Galician data are from Álvarez *et al.* (1986); for Balearic Catalan, see for example Veny (1982).

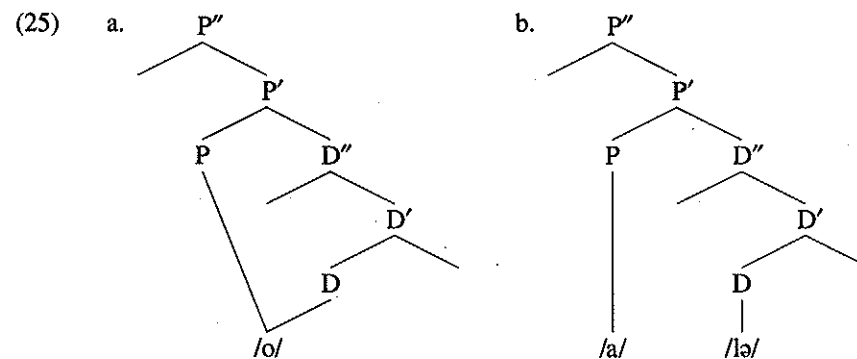
Spanish	P + pronoun 'with'	con	mí	conmigo
		con	tí	contigo
		con	sí	consigo
Balearic Catalan	P + Art(MASC) 'with'	əm(b)	əs	ən so
		əm(b)	əs(PL)	ən sos

The fact that in contraction two different lexical items are involved distinguishes it clearly from word internal allomorphy. Compare cases of contraction with cases of plain, internal allomorphy:

(24)	a.	b.	c.	d.	
	[ <sub>V</sub> êtes]	[ <sub>V</sub> [ <sub>R</sub> (être)] [ <sub>Aff</sub> 2SG]]	[ <sub>Z</sub> au]	[ <sub>P</sub> à] [ <sub>D</sub> le]	(French)
	a'. [ <sub>V</sub> sois]	b'. [ <sub>V</sub> [ <sub>R</sub> (ser)] [ <sub>Aff</sub> is]]	c'. [ <sub>Z</sub> al]	d'. [ <sub>P</sub> a] [ <sub>D</sub> el]	(Spanish)

Inflectional verbal forms like the suppletive *êtes* '(you) are' (24a), or forms with an irregular root allomorph like *sois* (24a') '(you) are' can have their nonregular morphological structure generated in the lexicon: their internal constituents, shown in (24b, b') are sublexical, and the whole form is a lexical category  $V(=X^0)$ . A contraction like French *au* (24c) corresponds to two lexical categories, as shown by the noncontracted forms in (24d); the constituent Z (alternatively, the domain  $P \cap D$ ) cannot be assigned to any lexical category or projection. Similarly, in the case of Spanish *al* (24c'), Z contains the lexical categories P and D, but these cannot be assigned to phonological subparts of Z. In both cases the problem lies in the fact that, if allomorphy is lexical, there is no way to assign the allomorphic contracted forms to a lexical category. If P and D are assigned their noncontracted usual forms in the lexicon, then phonological rules should become unrestricted in the sense that they should be able to apply to specifically marked individual lexical items.

In other words, in a representation like (25a) there is no possible node Z solely dominating P, D, that can be independently justified, and in (25b) there is no possible phonological rule that will yield [o] from /a lə/.



The solution I propose in order to account for these cases is based on the premise, which I think is basically correct, that contractions do not yield to a syntactic or to a phonological explanation, but are derived from properties of lexical representations.

A lexical item can be thought of as a set of syntactic, morphological, semantic and phonological properties. If a phonological representation  $\Phi$  belongs to a lexical item  $\Lambda$ , we say that  $\Phi$  is the (lexical) *phonological interpretation* of  $\Lambda$ .  $L$  will designate the set of all lexical properties of  $\Lambda$ , excluding  $\Phi$ . It is well known that, at the morphological level, a morpheme can receive different lexical phonological representations (that is, allomorphs). It is the case, on the contrary, that a lexical item, that is, a (possibly derived or inflected) word, has normally only a single lexical phonological interpretation. Choice of allomorphs is a highly idiosyncratic phenomenon which is therefore restricted to the lexicon. Outside the lexicon the resolution of allomorphy is only exceptional.

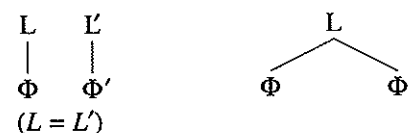
Let us review the different possibilities of word external variation, that is the possible cases in which there is not a one to one correspondence between  $\Phi$  and  $L$ . One case is free variation. A lexical item has two (or more) phonological interpretations which can appear freely. This is illustrated in (26a) with the Catalan word for 'nothing' which in many varieties can be either [re] or [res]. No allomorphic choice is imposed by the grammar, and both forms appear freely, independently of the context. Another case is phonologically controlled external allomorphy, discussed in section 2. In such cases it is possible to get a choice of the allomorph outside the lexicon, given exceptional conditions of emergency of the unmarked. (26b) illustrates this case with the choice between *en* and *el* in Catalan discussed in section 4. Finally, (26c) corresponds to the French contraction *au* = *à le*.

(26) Lexical categories with several phonological interpretations:

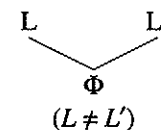
a. Free variation	b. External allomorphy	c. Contraction
re res 'nothing'	[ə] Juli     [l] Emili article        article	[o] ← [a] [lə] 'to the'

Since the choice between [o] and [a][lə] cannot be determined by the phonology or by the syntax, it has to be encoded in the lexical representation. The proposed lexical representations for (26a-c) are (27a-c), respectively. (27d) shows the normal case, one to one correspondence of  $L$  with a single phonological interpretation  $\Phi$ .

(27) a. Free variation     b. External allomorphy, phonologically controlled



c. Contraction     d. Normal case



$\Phi, \Phi'$ : Lexical phonological representation

$L, L'$ : Rest of lexical representation

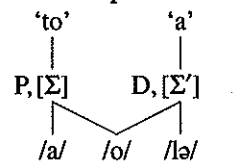
In the case of free variation (25a), the availability of two different, free phonological forms is due to the existence of two different lexical items which differ just in phonological shape. Since they are different lexical items, [re] and [res] will not compete for evaluation. Instances of phonologically conditioned external allomorphy conform to a lexical representation such as (27b): a single lexical item with two possible phonological interpretations,  $\Phi$  and  $\Phi'$ . These are both inputs to Gen and the total output is then submitted to optimal evaluation (see section 3). Contractions have the representation (27c), namely a representation in which two lexical items can be (exceptionally) interpreted by the same phonological representation (N.B., not by two different, identical phonological representations, as in free variation). Whereas at lexical insertion free variation (27a) presupposes a choice between two lexical items, and allomorphy (27b) forces two parallel interpretations (up to evaluation), contraction (27c) forces, under appropriate conditions, a single interpretation for two different lexical items.

Although this general picture has different technical implementations, I will illustrate the way it works assuming that lexical insertion proceeds as follows. The X-bar structure projected from the lexicon contains only those lexical properties that are relevant for syntax. At the point at which expressions are

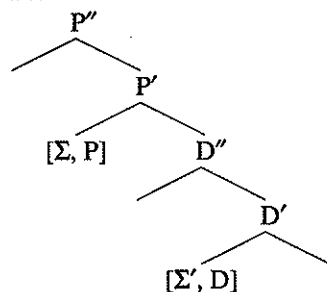
phonetically interpreted (Spell-out, Chomsky 1993), phonological properties are projected from the lexicon. Identity of properties belonging to the same lexical item is preserved (for example by co-indexing). The same should apply to properties relevant to LF.

Consider now the configuration of the relevant lexical and syntactic representations in a case of contraction:

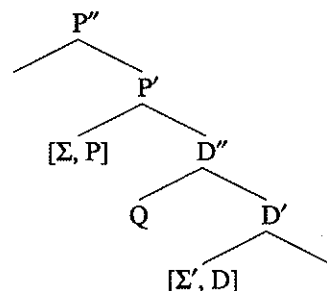
(28) a. French *à + le*,  
lexical representation



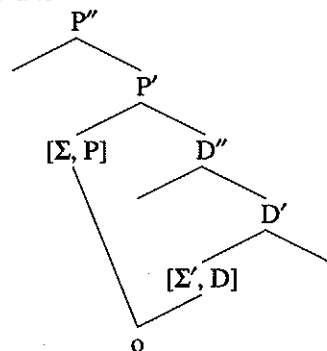
b. *à le*



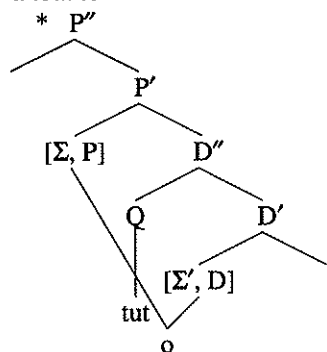
c. *à tout le*



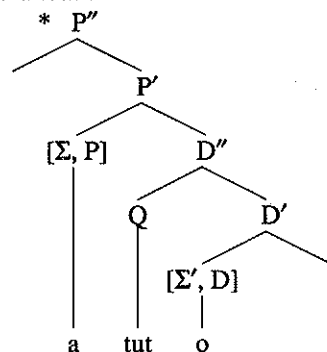
b'. *à le*



c'. *à tout le*



c''. *à tout le*



The lexical representation (28a) contains two lexical items; each one has two phonological interpretations, one of which is shared by both. This shared representation (/o/) behaves as the special case with respect to the Elsewhere Condition. The preposition and the determiner together with all other relevant syntactic properties  $\Sigma$ ,  $\Sigma'$ , are projected into X-bar structure (28b), ...*à le...*, and (28c), ...*à tout le...* (28c) differs from (28b) in the presence of lexical material intervening between P and D (notice that the presence or absence of intervening material is determined in the *derived* syntactic structure, that is at the point of Spell-out). At Spell-out, when linguistic expressions have to receive phonetic interpretation, lexical phonological properties are projected into the structures (28b, c), yielding (28b', c', c''). By the Elsewhere Condition the specific form /o/ takes precedence; it yields an acceptable structure if inserted into (28b), that is (28b'), ...*au...* When inserted in (28c), however, a crossing violation arises if *au* (/o/) is inserted: it both precedes and is preceded by *tout* (/tut/), a phonetically uninterpretable structure. Hence the elsewhere case is chosen and, *à* and *le* are inserted, giving rise to (28c'), *à tout le*.

Universitat Autònoma de Barcelona

References

Álvarez, Rosario, Henrique Monteagudo and Xosé Luis Regueira (1984). *Gramática Galega*. Vigo: Galaxia.

Cabré, Teresa (1993). Estructura Gramatical i lexicó: El mot mínim català. Unpublished Ph.D. dissertation, Universitat Autònoma de Barcelona.

Cabré, Teresa and Michael Kenstowicz (1995). Prosodic trapping in Catalan. *Linguistic Inquiry* 26: 695-705.

Chomsky, Noam (1993). A minimalist program for linguistic theory. In *A View from Building 20*, Kenneth Hale and Samuel Jay Keyser (eds.), 1-52. Cambridge, MA: MIT Press.

Hayes, Bruce (1990). Precompiled phrasal phonology. In *The Phonology-Syntax Connection*, Sharon Inkelas and Draga Zec (eds.), 85-108. Chicago: The University of Chicago Press.

McCarty, John and Alan Prince (1994). The emergence of the unmarked. *Proceedings of the North Eastern Linguistic Society* 24, vol. 2, 333-379. Amherst, MA: GLSA.

Piera, Carlos (1985). On the representation of higher order complex words. In *Selected Papers from the XIIIth Linguistic Symposium on Romance Languages*, Larry D. King and Catherine A. Maley (eds.), 287-313. Amsterdam: John Benjamins.

Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint Interaction and Satisfaction*. Unpublished manuscript, Rutgers University and University of Colorado.

Rohlf, Gerhard (1966). *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi. [Original German edition: Bern: A. Francke, 1949.]

Sistac, Ramon (1993). *El Ribagorçà a l'Alta Llitera*. Barcelona: Institut d'Estudis Catalans.

Tranel, Bernard (1994). French liaison and elision revisited: A unified account within optimality theory. Unpublished manuscript, University of California, Irvine.

Veny, Joan (1982). *Els parlars catalans*. Palma de Mallorca: Editorial Moll.