

CURRENT TRENDS IN PHONOLOGY MODELS AND METHODS

Edited by

Jacques Durand

and

Bernard Laks

With the support of:

CNRS, ESRI, University of Aix-Marseille,
University of Manchester, University of
Oslo, Paris X

First published 1996

Copyright © of chapters by individual authors

Copyright © of other material by the editors

Published by the European Studies Research Institute (ESRI),
University of Salford, Salford, Manchester M5 4WT

Printed in Great Britain by Rank Xerox

ISBN for complete set of 2 volumes: 1 901471 00 4

ISBN for this volume: 1 901471 02 0

External allomorphy as Emergence of the Unmarked

Joan Mascaró

Universitat Autònoma de Barcelona

1 Introduction

This paper is concerned with a long-standing theoretical and empirical problem. It is a theoretical problem because it regards the organization of the grammar; in particular it addresses the question of where in the grammar are lexical, unpredictable morpheme alternations to be included, and where are phonological regularities to be expressed. It is empirical in the sense that it has to do with a well-defined descriptive phenomenon, commonly referred to as *external allomorphy*, or *phrasal allomorphy*. External allomorphy is allomorphic, lexical variants (hence not determined by phonological processes) whose choice is conditioned outside of the normal domain of allomorphy, the word. The problem leads to two successive paradoxes, that are, we claim, solvable if we appeal to the notion of *emergence of the unmarked* of OT.

2. An illustrative example

Let us first consider a well known case of external allomorphy as a first illustration of the problem. The English indefinite article shows two phonetic forms [ən], [ə], that are not phonologically general (i.e. they are restricted to the lexical item *definite article*), hence not phonologically predictable, as (1a,b,c) show.

- | | | | | |
|-----|----|--------------------|-----------------|----------------------|
| (1) | a. | ə[n] | impossible word | a possible word |
| | b. | *[t]ə[n] | impossible word | the possible word |
| | c. | in impossible word | | *[i] impossible word |

This suggests an allomorphic, i.e. a lexical solution: /ə/, /ən/ are not retractable to a single phonological underlying form, but are both listed in the lexicon under the lexical item corresponding to the indefinite article. In other words, the lexical item *definite article* has not a single underlying phonological representation, but two. Now compare this

situation to a normal case of allomorphy, i.e. *internal* allomorphy like (2a):

- (2) a. make made meɪk meɪ + d
 say said seɪ se + d
 stand stood stænd stu + d
- b. (She) (αsai+d)α ((βan)β(impossible))β (word).

In (2a) the domain is the word, a lexical element (in its inflected form), whereas in (1) the domain of allomorphy includes more than one word—more exactly, it corresponds to some prosodic constituent that dominates the prosodic word. The sentence in (2b) contains two domains of allomorphy, a word domain α and a higher prosodic domain β. Since allomorphy is idiosyncratic, it is lexical. That would allow for a lexical solution for the word case (α), since words are lexical elements. But it raises a paradox for the external case (β): the alternation is *lexical*, but it takes place in a *nonlexical* domain.

The second problem appears once the conditioning of the alternation is examined more closely. Recall that cases of external allomorphy like English *alan* consist of two parts, the allomorphic alternants, i.e. /a/-/an/, and the external conditioning. This external conditioning is not like the alternation itself, which is idiosyncratic, unmotivated. The fact that [ən] appears before a vowel, and [a] before a consonant in (1) is not a coincidence, it is completely regular. As we will see (section 4), this kind of regularity is not confined to the particular case of *alan* allomorphy, but it extends homogeneously across languages.

Using the informally the notation reserved for derivational processes, we might say that we are confronted with a regularity in (part of) the *structural description*, and an idiosyncrasy in the *structural change*: the irregular, lexically listed forms *a*, *an* alternate (*alternate* is represented by ↔ in (3)), in the regular environments /__C and /__V, respectively:

- (3) a. a / __C ↔ an / __V

Summing up, we are faced with the following paradoxes:

- (4) a. Why is external allomorphy *lexically* restricted (e.g. restricted to the indefinite article), but regularly defined in a *nonlexical* context?
 b. Why is the context of allomorphic choice phonologically regular and at the same time undervivable by (post)lexical phonological rules?

3. External allomorphy and emergence of the unmarked

In this section it will be shown that the paradoxes in (4) disappear under the appropriate theoretical assumptions. For the moment I will assume that the features assigned to English *alan* allomorphy apply also to cases found in other languages; empirical evidence in that direction will be provided in the next section. The theoretical conditions under which a proper analysis of external allomorphy can be developed are provided by some properties of OT. The basic idea is first developed in McCarthy & Prince (1994), in their explication of some crucial properties of reduplication. Consider their analysis of Nootka as an illustration. I reproduce partly their example of Nootka (14) as (5) (I italicize the reduplicant):

- (5) ?u- ?u- 'ih 'hunting it'
 ʔi- ʔims- 'ih 'hunting bear'

The syllable structure of Nootka allows codas like [h] in (5), which means that PARSE-SEG and FILL dominate NO-CODA (constraints are defined below in (7)). The crucial point, however, is that "there is a particular class of syllables that cannot have a coda: syllables in the reduplicative formation." (McCarthy & Prince 1994: 345). This apparent contradiction dissolves once we realize that faithfulness constraints like PARSE and FILL have a special property, stated in (6): they always evaluate an output of Gen with respect to another form, the underlying form. If there is no underlying form to refer to, a faithfulness constraint cannot be applied, and therefore it cannot be violated.

- (6) *Faithfulness constraints:*

Given an input to Gen which is a lexical entry *ini*, and a candidate *candj* (*candj* ∈ Gen (*ini*)), satisfaction or violation(s) of a constraint belonging to the set of faithfulness constraints is a function both of *candj* and *ini*.

If faithfulness conditions cannot be violated when there is no underlying form, other constraints which are ranked lower in the hierarchy, and which normally show no effect because they are overpowered by the former, will *emerge* as deciding constraints. One such case, as argued in more detail in McCarthy & Prince (1994), is

reduplication, where *faithfulness* of the reduplicant to the base is controlled by MAX (7), a constraint which requires that every element of the base has a correspondent element in the reduplicant. For ease of reference I also list in (7) other constraints that will be considered in this paper (for further details, see McCarthy & Prince (1994), Prince & Smolensky (1993)).

- (7)
- 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.
 MAX: Every element of B (the base) has a correspondent in R (the reduplicant).

In the case of Nootka, the ordering NO-CODA >> MAX determines the choice. Even if the reduplicants *tjims* (8a) and *tjim* (8b) satisfy MAX better than *tji* (8c), because they diverge less from the base *tjims*, the coda-less CV form *tji* is preferred as the reduplicant because it satisfies better the higher ranked constraint NO-CODA:

(8)

	NO-CODA	MAX
a. <i>tjims-tjims-</i>	**!	
b. <i>tjim-tjims-</i>	**!	*
c. <i>tji-tjims-</i>	*	**!

We can now return to the case of external allomorphy. Assume, for simplification, that we have exactly two allomorphs, /A/ and /B/, and that we divide the set of constraints Con in two subsets Con₁ and Con₂, where no constraint in Con₂ dominates any constraint in Con₁. If some output of Gen (Gen (A/ ∩ /B)) is rated as the most harmonic by Con₁, then it will belong either to Gen (A/) or to Gen (/B/). If, say, it belongs to Gen (/A/), then /A/ will be the allomorph chosen, independently of the evaluation of the lower ranked set of constraints Con₂. If this is not the case, i.e. if there are two candidates *cand1* ∈ Gen (A/) and *cand2* ∈ Gen (/B/) that are equally harmonic with respect to Con₁, then Con₂ will be allowed to decide between them.

In our example if we take *cand1* = [a] and *cand2* = [ən] (as in, e.g., [a.sm], [ə.n], [ən.sm], [ən.n]), which show no empty (□) or unparsed (<X>) elements added by Gen, both *cand1* and *cand2* satisfy Con₁, i.e.

(among others) faithfulness constraints. But they do not show the same harmonic rating with respect to constraints that relate to syllabic structure. Lower ranked constraints in Con₂ will now *emerge* as decisive constraints and will determine the optimal candidate.

Both in the case of reduplication and in the case of external allomorphy, the fact that lower ranked constraints *emerge* as crucial constraints is due to the fact that the multiplicity of forms to be evaluated does not arise from the effect of Gen only, but also from other sources. In the case of reduplication the additional source is the process of copying from the base (or of the function Gen applied to the reduplicative morpheme RED—see McCarthy & Prince (1994:340-343)). In external allomorphy the source is the multiplicity of underlying forms in a single lexical item.

In the case of English, both /a/ and /an/ satisfy faithfulness constraints. After Gen has introduced syllable constituency to Art X sequences like *a(n) impossible word*, *a(n) possible word*, we get, among others, the forms [ə.n] *impossible word* and [ə .l] *impossible word* in the first case, and [ən .p] *possible word*, [ə .p] *possible word* in the second case. Faithfulness (and other) constraints being equally satisfied by both candidates in each pair, the burden of choice rests on ONSET and NO-CODA, which filter out the more marked syllabic structures VV ([ə V]) and CC ([ən C]):²

(9)

	ONS	NO-CODA
a.		
<i>ə.n impossible</i>	*	
<i>ə .l impossible</i>	**!	
b.		
<i>ən .possible</i>	*	*!
<i>ə .p possible</i>	*	

4. Three other cases

Cases like English /a-/an/ are by no means isolated. Although the existence of external allomorphy in a given language is limited (probably because of general properties of the structure of the lexicon), the phenomenon is quite widespread across languages. We will consider three other illustrative cases. See Mascaró (in press) for more examples.

In Moroccan Arabic the 3rd masc. sg. pronominal enclitic 'him, his' presents two allomorphs, [h]/[u]. Pronominal enclitics appear as objects after verbs, as obliques after prepositions and particles, and as obliques with a genitive reading after nouns. The pronominal enclitic of first person 'me, mine' is also subject to allomorphy; in this case it alternates as [i]/[ja]. The examples in (10), from Harrell (1962), show the differences in allomorphic form of these enclitics when they follow the lexical elements [xt'a] 'error', [ktab] 'book', [mʃa] 'with', [menn] 'from', [ʃafu] 'they saw', [ʃaf] 'he saw'. (10a,c) shows the clitic hosts ending in V, and (10b,d) those ending in C:

- (10) Moroccan Arabic: 3rd masc. sg. and 1st sg. clitic allomorphs
- | | |
|--------------------------|--------------------------|
| a. 3rd masc. sg. / ___ V | b. 3rd masc. sg. / ___ C |
| [xt'a h] 'his error' | [ktab u] 'his book' |
| [mʃa h] 'with him' | [menn u] 'from him' |
| [ʃafu h] 'they saw him' | [ʃaf u] 'he saw him' |
| c. 1st sg. / ___ V | d. 1st sg. / ___ C |
| [xt'a ja] 'my error' | [ktab i] 'my book' |
| [mʃa ja] 'with me' | [menn i] 'from me' |
- (No verbal examples because 'me' is [ni] after verbs.)

The allomorph is chosen depending on the phonological shape of an adjacent word (within a certain prosodic domain). In the 3rd masc. sg. pronominal the form [h] appears after vowels, and [u] appears after consonants. In the 1st sg., after a preceding vowel we get [ja], and after a preceding consonant we get [i]. Here again, the generalization is clear: the allomorphic lexical choice is determined by the degree of markedness, in syllabic structure terms, of the resulting combination: the less marked structure is chosen. In the case of the third person clitic, and when the host is vowel final, a structure with a closed syllable (C...Vh) is preferred to a structure (C...V)(Cui), i.e. a closed syllable fares worse than an onsetless syllable. In the case of consonant final hosts, a CV structure like (C...Ci) is preferred to the structure (C...Ch), in which we get a syllable closed by two consonants.

For the 1st sg. enclitic, the situation is parallel: (C...V)(Cja) with a ...V.CV structure is preferred to (C...V)(Cj) with a ...V.V structure. For the structure ...C.CV of (C...C)(Cja), which contains a closed syllable.

Within OT, we directly derive this descriptive generalization from the multiplicity of underlying phonological forms in the lexical representation of these clitics. We first consider the syllable structure of Moroccan Arabic: multiple onsets, onsetless syllables, and multiple codas are possible (cf. Kiebt 'I wrote, *aji* 'come!-sg.' (Harrell 1962: 42, 43). Hence faithfulness constraints must overrank the syllable structure constraints ONSEF and NO-CODA. If, as dictated by OT, these two constraints *must* be part of the grammar, and assuming the normal order ONSEF >> NO-CODA, then whenever faithfulness constraints are not applicable, the syllable structure constraints will determine the allomorphic choice. In the case of *men.n ja* / *men.n i* we also need the constraint COMPLEX so that the less complex onset is chosen (alternatively, depending on syllabification, the less complex coda in *men.n ja* over the more complex one.

- (11)
- | | | | |
|----|-----------|---------|---------|
| a. | ONS | NO-CODA | COMPLEX |
| ⵜⴰ | xt'a h | | * |
| ⵜⴰ | xt'a. u | *! | * |
| b. | ONS | NO-CODA | COMPLEX |
| ⵜⴰ | ktab h | | *! |
| ⵜⴰ | ktab. b u | | * |
| c. | ONS | NO-CODA | COMPLEX |
| ⵜⴰ | mʃa. ja | | * |
| ⵜⴰ | mʃa. i | *! | * |
| d. | ONS | NO-CODA | COMPLEX |
| ⵜⴰ | men.n ja | | * |
| ⵜⴰ | men.n i | * | *! |

Another well known case of external allomorphy is the alternance found in French words like, *nouvel-nouveau, vieil-vieux, fol-fou, cet-cé, mon-ma, ton-tu, son-sa*. When a following word begins with a vowel, the

first member of the pair is chosen: if it begins with a consonant, then the second member is chosen.³

(12) French Belle allomorphy: *beau* ~ *bel*, *nouveau* ~ *nouvel*, etc.

- a. / ___ V
- | | | |
|-----------|-------------|-------------|
| bel ami | beau mari | il est beau |
| [bel ami] | [bo mari] | [bo] |
| | beau à voir | [bo] |
- b. / ___ V
- | | | |
|---------|---------|---------|
| nouvel | nouveau | nouveau |
| [nuvel] | [nuvo] | [nuvo] |
| vieil | vieux | vieux |
| [vʁeɪ] | [vʁø] | [vʁø] |
- c. joli ami
- | | | |
|---------------|----------------|-----------|
| deux [z] amis | joli mari | joli |
| quel ami | deux [Ø] maris | deux [Ø] |
| petit [t] ani | quel mari | quel |
| | petit [Ø] mari | petit [Ø] |

Belle allomorphy is similar to the English case presented in the second section. Faithfulness constraints being satisfied by both candidates, the candidate that succeeds in obtaining a better syllabification with the following word will be the optimal candidate. Again, ONS and NO-CODA are the crucial constraints. Notice that in this case the word subject to allomorphy can be clearly domain final in *beau*, *nouveau*, *vieux*, and *fou*. (cf. fn. 3).

(13)

	ONS	NO-CODA									
a.	<table border="1"> <tr> <td>ɛə</td> <td>be.l.a.mi</td> <td></td> </tr> <tr> <td>ɛə</td> <td>bo .a.mi</td> <td>*!</td> </tr> </table>	ɛə	be.l.a.mi		ɛə	bo .a.mi	*!				
ɛə	be.l.a.mi										
ɛə	bo .a.mi	*!									
b.	<table border="1"> <tr> <td></td> <td>ONS</td> <td>NO-CODA</td> </tr> <tr> <td>ɛə</td> <td>bel .ma.ri</td> <td>*!</td> </tr> <tr> <td>ɛə</td> <td>bo .ma.ri</td> <td></td> </tr> </table>		ONS	NO-CODA	ɛə	bel .ma.ri	*!	ɛə	bo .ma.ri		
	ONS	NO-CODA									
ɛə	bel .ma.ri	*!									
ɛə	bo .ma.ri										
c.	<table border="1"> <tr> <td></td> <td>ONS</td> <td>NO-CODA</td> </tr> <tr> <td>ɛə</td> <td>bel</td> <td>*!</td> </tr> <tr> <td>ɛə</td> <td>bo</td> <td></td> </tr> </table>		ONS	NO-CODA	ɛə	bel	*!	ɛə	bo		
	ONS	NO-CODA									
ɛə	bel	*!									
ɛə	bo										

I will finally present a case in which constraints which are higher ranked than ONS and NO-CODA play a role in the evaluation.

In Catalan there are two types of definite articles. The common form is *el/la/els/les*. The so-called *personal* article is used with (unique) proper personal names and has (in Central Catalan) two forms for the masc. sg. One is *en*, the other is identical to the corresponding form of the definite article, /, phonetically [ən] and [l], respectively. Thus whereas *en Wittgenstein* has only a unique interpretation, and *el Wittgenstein* only a nonunique one (as in *el primer Wittgenstein* 'the first Wittgenstein'), *l'Einstein* has both interpretations. This is the reason why the NPs in (14a) cannot pluralize, while the nonunique NPs in (14b) can: *he llegit els Prince* (1976 i 198), *els dos Wittgensteins* 'the two W.', *els Einsteins que desconexem* 'the E. we don't know'.

- (14) a. *Definite personal*
- | | |
|-----------------|-----------------------------|
| en Prince | He llegit el Prince (1976) |
| l'Alan Prince | l' read' |
| en Wittgenstein | el (primer) Wittgenstein |
| l'Einstein | l'Einstein (que desconexem) |
- b. *Definite nonpersonal*
- | | |
|--|-----------------|
| | we didn't know' |
|--|-----------------|

Thus the personal definite article has two allomorphs, /ən/ and /l/, the latter coinciding formally with the morph of the nonpersonal definite article. The choice between the two allomorphs is determined by the shape of the following word, in parallel to the cases examined above: a following vowel causes the appearance of /l/, and a following consonant the appearance of /ən/. It should be noted that #lC as it stands (as in [l

prins]) would violate the constraints regulating possible onsets: in such cases insertion takes place, and violation of FILL is circumvented by satisfaction of those constraints that control sonority sequencing in onsets. Therefore the form with *e* to be compared to the form with *en* in (15) is supplied with *ɲ*, which will show up as an epenthetic schwa:

(15)

	FILL	ONS	NO CODA
an .prins		*	*
ɲ .prins	*ɲ	*	*
a.n a.lan prins		**ɲ	**ɲ
ɲ a.lan prins			*

In the case of the definite nonpersonal article (as in *He lleġgji [al prins]* (1975)), there is no allomorphy. A single underlying form /l/ will show up as [l] before vowels, and as [ɲl] before consonants. We are therefore faced in this case with a normal instance of multiplicity of forms arising only from applying Gen to a single underlying form.

5. Summary and conclusion

Two basic properties of OT are that constraints are ranked, and that underlying forms are submitted to Gen, a general function providing a set of possible alternative analyses to its inputs. Given a lexical item L_i with a phonological form ϕ_i , we normally get $Gen(\phi_i) = \{cand_1, cand_2, \dots, cand_n\}$ as the output of Gen. This output is then evaluated by the set of constraints, yielding a single, optimal candidate, $cand_j = Eval(\{cand_1, cand_2, \dots, cand_n\})$. But under special conditions the output of Gen is not the result of a single underlying form. Under such circumstances a subset of higher ranked constraints fails to decide on the optimal output, i.e. it yields the same harmonic rating to two (or more) candidates. This set of tied candidates will be differently rated by lower ranked constraints which will thus emerge as crucial for the evaluation. External allomorphy, like reduplication, is an instance of such special conditions. The existence of multiple underlying phonological forms for one lexical item can result in ties of harmonic ratings with respect to faithfulness constraints for several candidates, since each allomorph can be equally faithful to its own underlying form. When the outputs of Gen are evaluated with respect to constraints that are lower in the hierarchy, these

turn out to give a different rating of these candidates. In the examples examined in this paper, OT correctly predicts that the allomorph whose syllabification with an adjacent word results in a less marked syllabic structure will be favored over the rest, and chosen by Eval as the optimal candidate.

NOTES

- [1] I will not address here the problem of determining the domain of external allomorphy, surely an important issue.
- [2] The analysis predicts that the definite article, when appearing with no context to the right (i.e., when final in the prosodic domain), should take the form *a*, and not *an*, the latter violating NO-CODA. Although such contexts are not very common, cf. sentences like (i) and (ii):
 - (i) I was talking about A, I was not talking about THE optimal candidate.
 - (ii) I was talking about *AN, I was not talking about THE optimal candidate.
- [3] For a discussion of the analysis in Tranel (1994), see Mascaro (in press).

REFERENCES

- Harrell, R. S. (1962) *A short reference grammar of Moroccan Arabic*. Washington D.C.: Georgetown University Press.
- Hayes, B. (1990). Precompiled phrasal phonology. In S. Inkelas and D. Zec (eds.) *The Phonology-Syntax connection*. Chicago: University of Chicago Press. 85-108.
- McCarthy, J. & A. Prince (1994). The emergence of the unmarked. Optimality in Prosodic Morphology. *Proceedings of the North Eastern Linguistic Society*, vol. II. Amherst: University of Massachusetts. 333-379.
- Mascaro, J. (in press). External allomorphy in Romance. *Probus*.
- Piera, C. (1985). On the representation of higher order complex words. In L. D. King and C. A. Maley (eds.) *Selected Papers from the XIIIth Linguistic Symposium on Romance Languages*. Amsterdam/Philadelphia: John Benjamins. 287-313.
- Prince, A & P. Smolensky (1993). *Optimality Theory: constraint interaction and satisfaction*. University of Massachusetts and University of Colorado. Ms.
- Tranel, B. (1994). French liaison and elision revisited: a unified account within optimality theory. UC Irvine. Ms.