

Valencian Hypocoristics: When Morphology Meets Phonology

Eduard Artés

CLT - Universitat Autònoma de Barcelona

1 Introduction

Vowel epenthesis is a crosslinguistically widespread strategy to satisfy language particular structural requirements (Hall 2011). An example of this can be found in Peninsular Spanish:

- (1) /stóp/ > [estóp] A tautosyllabic sequence of two obstruents is forbidden, so vowel epenthesis applies in order to split the cluster into two syllables.

A different example of epenthesis is the so called ‘morphological epenthesis’ (Cardinaletti & Repetti 2008), in which a vowel with morphological content is used to solve a phonological problem:

- (2) *l'albero* ‘the tree’ [l#V]
lo spettacolo ‘the show’ [lo#sC]

Assuming the underlying representation of the Italian masculine definite article to be /l/ (Artés 2013), the masculine exponent *-o* is inserted when needed for syllabification. Allomorphy lies on inflection (*-Ø/-o* for masculine) and there is no need to argue for allomorphy in the lexical entry ([l]/[lo]). The insertion of an epenthetic vowel (*i* in Italian) would occupy a morphological position (/l-MASC.SG./), specifically the gender space, so the use of a morphologically meaningful vowel is preferred.

The premises of this article will be the following:

1. There is avoidance of epenthesis as far as possible: morphs matching morphosyntactic features are preferred as a repair strategy. In case of phonological conflict, the preservation of structure already present in the grammar (i.e., gender exponents) will be favored.
2. Epenthetic elements cannot occupy inflectional positions. The mismatch between morphological and phonological structure motivates the use of other gender morphs when a phonological problem arises.

One of the goals of this paper is to elucidate the mechanisms that determine gender exponence. This will be developed in section 2. Section 3 sketches the truncation patterns of Catalan and the creation of Valencian hypocoristics. An OT analysis is shown in section 5, while sections 6 and 7 present the theoretical implications of the proposal and the conclusions, respectively.

2 Allomorphy in inflection

Catalan is a rich inflectional language with a binary system in gender and number (masculine-feminine, singular-plural). As a Romance language, the right edge of the word is the placement for inflection (see Table 1).

* I am very grateful to Eulàlia Bonet, Donca Steriade and Francesc Torres-Tamarit for their valuable comments on previous versions of this proposal. I would also like to thank Adam Albright, Paul de Lacy, Edward Flemming, Michael Kenstowicz and Katya Pertsova for their suggestions. Any errors contained herein are mine. Research for this article was funded by grant FFI2010-22181-C03-01 from the Spanish *Ministerio de Economía y Competitividad*.

	Unmarked	<i>example</i>	Marked	<i>example</i>
Masculine	Ø	<i>cor</i> [Ø] ‘heart’	u>ə	<i>llor</i> [u] ‘parrot’, <i>par</i> [ə] ‘father’
Feminine	ə	<i>cas</i> [ə] ‘house’	Ø	<i>mel</i> [Ø] ‘honey’

Table 1: Central Catalan nominal system (singular)

As shown above, there is a default exponent in the masculine and feminine forms and other exponents which are labeled as marked.¹ The vast majority of masculine nouns and adjectives have a zero exponent (due to historical apocope in Catalan), whereas the feminine is usually specified with an *-a* ending.

There are different theories of inflectional morphology within the generative tradition. I will assume here a root-driven approach in which the inflectional morphemes are underdetermined in the input (following Mascaró 2007).

The inflective allomorphs are listed in the lexicon together with a total ordering of the set. The ordering follows from the markedness hierarchy established in Table 1. Finally, the choice of the allomorphs takes place in the phonology, where they compete for insertion. The representation looks as follows:

- (3) Masculine: /kór{Ø>u>ə}/
Feminine: /káz{ə>Ø}/

For phonology to decide over the allomorphs, CON needs to provide a constraint that favors one of the alternatives (e.g., -Ø, -u or -ə for the masculine). A faithfulness constraint PRIORITY (Mascaró 2007) will do the job:

- (4) PRIORITY
Respect lexical priority (ordering) of allomorphs.
Given an input containing allomorphs m_1, m_2, \dots, m_n , and a candidate m_i , where m_i is in correspondence with m_i , PRIORITY assigns as many violation marks as the depth of ordering between m_i and the highest dominating morph(s).

PRIORITY, by definition, gives preference to the highest ranked allomorph, favoring the assignment of the unmarked exponent (-Ø for masculine, -a for feminine). However, as we have seen, other inflectional endings surface sometimes (e.g., *par*[ə], masc.sg.). Thus, we need another constraint that prevents the assignment of the default exponent in all cases:

- (5) RESPECT
Respect idiosyncratic lexical specifications (Bonet et al. 2007)

RESPECT is a constraint that tries to encode the lexical idiosyncrasies of languages. Any theory needs to deal with exceptions and phenomena that cannot be derived by general mechanisms, and RESPECT is an attempt to formalize this problem. Like English irregular plurals (e.g., *scheme-schemata*), the marked cases need to be somehow specified in the lexicon. A word like *pare* ‘father’ will then have the following representation:

- (6) /parə{Ø>u>ə}/

RESPECT, due to its specificity, will generally outrank PRIORITY. It is important to notice that these morphological constraints are interleaved with purely phonological ones. A clear example of it can be found in the third person masculine singular clitic /l/:

- (7) a. *l’agafo* ‘I take it’: /l#əyáfu/ > [ləyáfu]
b. *el troba* ‘She/He finds it’: /l#tróβə/ > [əltróβə]²
c. *trobar-lo* ‘to find it’: /truβár#l/ > [truβárlu]
d. *troba’!* ‘find it!’: /tróβə#l/ > [tróβəl]

¹ The division between marked and unmarked relies on the frequency of appearance of the inflectional endings.

² Throughout the article, the epenthetic segments will be shown in italics.

The syllabification of the clitic determines the insertion of an epenthetic schwa in (7b) (an onset cluster of the type liquid+stop is forbidden in Catalan), whereas in (7c), the illegal coda cluster is avoided via the insertion of the marked masculine allomorph *-u*. As already mentioned, this can be understood as a phenomenon of ‘morphological epenthesis’. Morphology as a repair strategy is preferred over the creation of new phonological material, a fact that resorts to ‘Morphological Recoverability’ (van Oostendorp 2006):

- (8) Morphological Recoverability: phonological structure mirrors morphological structure as closely as possible

The insertion of a schwa in (7c) would occupy the morphemic space of gender, so the grammar chooses to be conservative and recycles the exponents used in other contexts (but always referring to gender).

With these assumptions, we can now switch to the analysis of the hypocoristics.

3 Truncation processes in Catalan

Truncation is a morphological process that separates a sequence from the base. There are two basic patterns of truncation, namely subtractive and templatic. Subtractive truncation consists in the deletion of some phonological material from the stem while in templatic truncation the output form has to adapt to a prosodic template (Alber & Arndt-Lappe 2012). This is the case for Spanish (examples in (9)) or Catalan:

- (9) Umberto > Beto [béto]
 Ignacio > Nacho [nátʃo]
 syllabic trochees ('σσ), the basic foot of Spanish

Cabré (1994) has argued that nicknames respect the minimal word requirement of Catalan, i.e., words need to be minimally bimoraic. Catalan being a quantity sensitive language, this requirement creates monosyllabic or disyllabic sequences with a right-to-left alignment. In the disyllabic nickname, each vowel projects a mora and builds a bimoraic foot, whereas in the monosyllabic word the coda also bears a mora and gives rise to the same prosodic structure. There are four basic patterns (Cabré & Kenstowicz 1995):

CV pattern	Hypocoristic	Foot structure
'CV.CV	Agustina > Tina	(tí _μ nə _μ)
'CVC.CV	Vicenta > Centa	(sɛ́ _μ n _μ)tə
'CVC(C)	Isabel > Bel	(bɛ́ _μ l _μ)
CV.'CV(C)	Bartomeu > Tomeu	tu(mɛ́ _μ w _μ)

Table 2: Truncation patterns

Moreover, the empirical facts demonstrate templatic truncation is an output-to-output derivation (Benua 1995, Benua 1997), as can be shown in the following two examples:

- (10) (Isa)(bel) > (bel), *sa(bel)
 (Bar)to(meu) > (meu), to(meu)

The hypocoristics are created from an output with feet already built in it. If it were an input-output relationship, **Sabel* and *Tomeu* would be both perfectly possible. However, only *Tomeu* is a possible output, because the parsing of *(Isa)(bel)* doesn't allow the adjunction of the syllable *-sa*,³ violating a constraint PW-2σ that requires the hypocoristic to be two syllables long. In *(Bar)to(meu)*, by contrast, there are two well-formed feet that leave an unparsed syllable *-to*, which is attached to the hypocoristic to create a disyllabic form (the unmarked structure, which accomodates to the constraint PW-2σ). Cases like *meu* ‘are a recursive truncation from disyllabic hypocoristics’ (Cabré 1998).

³ As noted by Cabré (1998), there is a constraint PARSE-SYLL in Catalan that requires syllables to be parsed into feet. **Sabel* does not respect it as it implies the existence of unparsed syllables in the base *–*Isa(bel)–*.

3.1 Valencian hypocoristics Valencian⁴, a dialect of Western Catalan, has a different behavior in the domain of hypocoristics. They seem to be quantity insensitive, as the only acceptable outputs need to conform to a disyllabic pattern. Furthermore, they must be left-headed, too.

Name	Central Catalan	Valencian
Agustina	Tina	Tina
Francisco	Cisco	Cisco
Isabel	Bel	<i>Bela</i>
Miquel	Quel	<i>Quelo</i>
Manel	Nel	<i>Nelo</i>

Table 3: Differences between Central Catalan and Valencian

The question we have to address is why we need an overt gender marker instead of an epenthetic element (which would be *-e* in the case of Valencian). A possible option is that the hypocoristic requires the presence of gender inflection. This is not true after analyzing cases like *Sari*:

(11) Rosario > Sari

-i is not a feminine exponent in Catalan. This demonstrates that what is really needed is a vowel that projects a mora in order to create a bisyllabic trochee, be it a gender marker or just part of the root. Epenthesis is avoided in favor of the insertion of a morph, which maintains a better morphophonological correspondence. The assumptions of the relationship between morphological and phonological structure will be exposed in the following section.

4 A model of morphophonological correspondence

Following Walker & Feng (2004), I will assume in this paper a ‘Ternary Model of Morphology-Phonology Correspondence’. As quoted by the authors, “(...) this model extends to incorporate input-output mappings involving morphological structure, and it regulates mappings (i.e., indexing) between phonological and morphological structure in an output”. The idea is to extend the Correspondence Theory model (McCarthy & Prince 1995) to mappings between morphology and phonology.

There are three types of correspondence relations:

- morphology - morphology (input - output)
- phonology - phonology (input - output)
- *morphology - phonology (output - output)*

Inputs encode morphological and phonological information. Different types of constraints are sketched in Walker & Feng (2004)’s paper, but the ones crucial for the present analysis are those regulating the indexation between morphological and phonological outputs. In particular, it is important to determine whether a morpheme has a correspondent morph or how can an epenthetic element be linked (in case it can) to a morpheme. Instead of MAX and DEP, two other constraints are proposed:

- (12) MAX-MP: Every morpheme in the output is indexed with some phonological element in the output.
 MAX-PM: Every phonological element in the output is indexed with some morpheme in the output.

MAX-PM tries to give every morpheme a phonological counterpart. Thus, zero exponence would be a violation of MAX-MP. MAX-PM, by contrast, tries to avoid epenthetic elements, as every phonological element in the output must have a morphological correspondent (i.e., it must be indexed to a morpheme). An

⁴ Valencian, as opposed to Central Catalan, has a different system of vowel reduction. The allomorphs will thus be {Ø>o>e} for masculine and {a>Ø} for feminine.

epenthetic segment does not have any morphological filiation, incurring in a violation of MAX-PM. This is the case of Valencian hypocoristics:

- (13) Miquel
 > Quelo [kéllo] (o \Leftrightarrow MASC.SG.)
 > Quele [kéle] (epenthesis = not indexed)

Quele is not the winning candidate because the epenthetic *-e* is not indexed to the gender morpheme. It is outside the morphological structure:

- (14) ([kélØ]e)

Instead, *-o*, a masculine exponent used in other inflective contexts, is inserted. This is the strategy used by the grammar to avoid the creation of new phonological material. When possible, morphology is at rescue. Given these premises, we can now switch to the formal analysis in the next section.

5 The analysis


The analysis of the Valencian hypocoristics will follow a transderivational account (Benua 1997). Apart from MAX-PM and PRIORITY, these are the other constraints involved in the analysis:

- (15) MAXBT: every segment in the base has a correspondent in the truncated form (Benua 1995)
 HEADMATCH: if α is in $H'(\text{PrWd})$ and $\alpha R \beta$, then β is in $H'(\text{PrWd})$ (McCarthy 2000)
 SYLLTROCHEE: Feet must contain two syllables and be left-headed

MAXBT penalizes for every segment deletion in the hypocoristic. Together with SYLLTROCHEE, they accomodate the bimoraic pattern while maintaining the maximum resemblance to the base. HEADMATCH prevents the change of stress location, keeping the same foot head in the base and in the hypocoristic.

In (16) we can see the masculine hypocoristic formation:

(16)


Base: manélØ	HEADMATCH	SYLLTROCHEE	MAX-PM	PRIORITY	MAX-BT
a.  néllo				1	2
b. nélØ		1W		L	2
c. néle			1W	L	2
d. néle				2W	2
e. mánelØ	1W			L	L
f. manélØ		1W		L	L

HEADMATCH, SYLLTROCHEE, MAX-PM \gg PRIORITY, MAXBT

Candidate f) is eliminated due to the prosodic pattern requirement (an iamb is not permitted), whereas candidate e) is rejected because of the stress shift. *nélØ* doesn't respect the bisyllabic condition, so candidates a), c) and d) are the only possible outputs at this point of the evaluation. The adjunction of an epenthetic element violates MAX-PM, which is higher ranked than PRIORITY, making *néle* a bad candidate. The final decision is then left to PRIORITY, that has the *-o* allomorph in a higher position than *-e*. Thus, *nélo*, although violating PRIORITY, satisfies MAX-PM and becomes the winning output, as *néle* incurs in two violations of PRIORITY.

In the case of the feminine form, the analysis is practically the same, as shown in the following tableau:

(17)


Base: izaβélØ	HEADMATCH	SYLLTROCHEE	MAX-PM	PRIORITY	MAX-BT
a.  bél _a					3
b. bélØ		1W		1W	3
c. bélé			1W	1W	3
d. bélé				2W	3
e. ízaØ	1W			1W	3
f. zaβélØ		1W		1W	1L

PRIORITY inserts the default exponent in candidate a) because it already satisfies the bisyllabic condition, so there is no need to refer to lower ranked allomorphs as in the previous tableau.

We should also contemplate an option not manifested in the tableau. Hypothetically, we could have a candidate like *íza* in which the *-a* ending would correspond to a fusion of the root and the gender exponent (*íza-a* > *íza*). If that were the case, it would not violate PRIORITY, but it would be anyway discarded by a violation of HEADMATCH. Candidate a) is thus the only possible output.

Along this article it has been proposed that epenthesis cannot occupy morphological positions. The need to have morphological and phonological structures as similar as possible avoids the use of epenthesis as a repair strategy. Does that mean that epenthetic material never surfaces in Catalan? The answer is no, because it can appear when not occupying inflectional positions⁵. Tableau (18) shows how that can happen in initial positions, where inflection doesn't play any role.

(18)

/stóp{Ø>o>e}/	SONSEQ	CONTIGUITY	MAX-PM	PRIORITY
a.  estópØ			1	
b. estópo			1	1W
c. stópØ	1W		L	
d. setopØ		1W	1	

SONSEQ ≫ MAXPM

A high-ranked constraint SONSEQ is the responsible for the addition of epenthetic material, which rules out candidate c), the most faithful output. CONTIGUITY, against non-peripheral epenthesis, prevents candidate d) from winning, and determines *estóp* to be the winner. Epenthesis tries to be avoided, but at the left edge of the word there is no other possible option because the illegal onset cluster '*st*' needs to be syllabified somehow.

6 Theoretical Implications

The analysis presented so far relies on several assumptions from a theoretical point of view which, in turn, are reinforced by the empirical facts. It seems to advocate for a lexical-realizational model of inflectional morphology for several reasons:

1. The exponents need to be listed in the lexicon. Trying to explain the facts from an inferential perspective doesn't seem a possible option. Which morphological rule can we establish that determines the insertion of the marked allomorph in certain cases which are phonologically conditioned? The rule should be too specific and would lose explanatory power.

Moreover, other Catalan phenomena analyzed by Bonet et al. (2007) and Bonet (2008) within the same line of this article are equally better captured if we treat the exponents as independent units. In Catalan, the plural is created with the adjunction of an *-s* morph, as shown in the following example:

⁵ The focus of the proposal is the hypocoristic formation and, by extension, the nominal system. The analysis can be extended to the verbal morphology, but that escapes the aim of this article.


- (19) *cap* ‘head’ > *caps* ‘heads’

Certain masculine nouns and adjectives ending in a sibilant, though, need to insert a gender marker *-u*⁶ in order to avoid an OCP violation:

- (20) *gos* ‘dog’ > *goss[u]s* ‘dogs’⁷

Again, the explanation is straightforward if we assume the allomorphs to be lexically stored and only used when required by structural constraints (i.e., syllabification). The OT analysis proposed by Bonet (2008) looks similar to the one presented here (without the assumptions of Walker & Feng 2004’s proposal):

- (21)

/gós+{Ø>u}+s/		OCP	DEP	PRIORITY
a.	 gósus			*
b.	góss	*!		
c.	góssəs		*!	

PRIORITY is responsible for the choice of the marked allomorph. Having the exponents stored in the lexicon licenses its use in phonologically conditioned contexts, in this case a homophony avoidance (the higher ranked constraint OCP motivates the insertion of the morph *-u*, which breaks the sequence of sibilants). In this way, it is easily explainable how a phonological problem forces the use of other lexically stored allomorphs. The context is only available once the concatenation of morphemes has created the problem, so as previously exposed a lexical model seems to better adequate to the Catalan facts.

In Table 1 it was shown how certain nouns present a *-u* exponent in the masculine singular (e.g., *llor[u]* ‘parrot’). In this class of nouns, the formation of the plural yields a sequence like *llor[us]*, identical to *goss[us]*. An alternative explanation would be to assume that nouns ending in a sibilant in the singular are subcategorized (*à la* Paster 2006) to surface with a *-u* in the plural, but this idea should be rejected if we want to capture a linguistic generalization: it is the phonological conflict what motivates the insertion of the *-u* morph provided by morphology.

2. A realizational account also captures the facts better because it is the morphosyntactic features what license the insertion of the morphs:

MASC.SG. $\Leftrightarrow \{\emptyset > o > e\}$

FEM.SG. $\Leftrightarrow \{a > \emptyset\}$

The property MASC.SG., already active without the suffix (as opposed to incremental theories, cf. Stump 2001), is what determines the insertion of *-u* in the hypocoristic domain. Epenthesis cannot occupy the morphological space of these features, as they need to be filled with a matching morph (which will be finally chosen by constraint interaction).

Furthermore, a non trivial implication of the analysis is that allomorph selection is controlled by phonology. Several phenomena (the clitic combinations of section 2 and the OCP problems mentioned above) have proven to be cases of output optimization: well-formedness constraints impose the use of an allomorph that generates a better output. Additionally, the use of PRIORITY to select the correct allomorph –even in cases when there is no phonological conflict– implies that exponence is in part governed by CON.

⁶ NB: in Medieval Catalan, the inserted element was a schwa, the regular epenthetic segment of Catalan. The fact that the chosen vowel in modern Catalan is *-u*, the gender marker, reinforces the idea that there has been some kind of restructuring of the mapping between morphology and phonology to avoid mismatches in interpretation.

⁷ The two consecutive sibilants in *gossos* are just a matter of orthography. There is no gemination and it is pronounced as [s].

7 Conclusions

All in all, the data and the analysis presented so far point to a model in which phonology and morphology are closely entangled. Once phonological insertion is completed, morphology is still active and motivates the insertion of an inflectional exponent to accommodate to a prosodic pattern. PRIORITY is the powerful mechanism responsible for this, which interacts with markedness constraints that force the choice of a more marked allomorph in the hierarchy.

In addition, it has been shown that morphological space cannot be filled with epenthetic material. The grammar prefers to recycle the exponents listed in the lexicon, preventing the creation of new phonological material and maintaining a one-to-one correspondence between morphological and phonological structure (formalized through the MAX-PM constraint). Well-formedness conditions, however, may cause violations of this correspondence, which motivate the use of epenthetic segments, although these appear crucially in non-morphological positions only.

References

- Alber, Birgit & Sabine Arndt-Lappe (2012). Templatic and Subtractive Truncation. Trommer, Jochen (ed.), *The Phonology and Morphology of Exponence*, Oxford University Press.
- Artés, Eduard (2013). Morphological Epenthesis in Romance: a Case for Lexical Conservatism. *Poster presented at the 39th Incontro di Grammatica Generativa*.
- Benua, Laura (1995). Identity Effects in Morphological Truncation. *University of Massachusetts occasional papers in linguistics* 18, 77–136.
- Benua, Laura (1997). *Phonological Relations Between Words*. Ph.D. thesis, University of Massachusetts Amherst.
- Bonet, Eulàlia (2008). Item-and-Arrangement or Item-and-Process? *Cuadernos de Lingüística* XV, 1–12.
- Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró (2007). Allomorph Selection and Lexical Preferences: Two Case Studies. *Lingua* 117:6, 903–927.
- Cabré, Teresa (1994). Minimality in the Catalan Truncation Processes. *Catalan Working Papers in Linguistics* 4, 1–22.
- Cabré, Teresa (1998). Faithfulness to Prosodic Edges. Dialectal Variation in Truncated Words in Catalan. *Catalan Working Papers in Linguistics* 6, 7–22.
- Cabré, Teresa & Michael Kenstowicz (1995). Prosodic Trapping in Catalan. *Linguistic Inquiry* 26:4, 694–705.
- Cardinaletti, Anna & Lori Repetti (2008). The Phonology and Syntax of Preverbal and Postverbal Subject Clitics in Northern Italian Dialects. *Linguistic Inquiry* 39:4, 523–563.
- Hall, Nancy (2011). Vowel Epenthesis. Marc Van Oostendorp, Elizabet Hume, Colin J. Ewen & Keren Rice (eds.), *The Blackwell Companion to Phonology*, Oxford: Wiley-Blackwell.
- Mascaró, Joan (2007). External Allomorphy and Lexical Representation. *Linguistic Inquiry* 38:4, 715–735.
- McCarthy, John J. (2000). The Prosody of Phase in Rotuman. *Natural Language & Linguistic Theory* 18:1, 147–197.
- McCarthy, John J. & Alan Prince (1995). Faithfulness and reduplicative identity. *University of Massachusetts Occasional Papers in Linguistics* 18.
- van Oostendorp, Marc (2006). A Theory of Morphosyntactic Colours. *Ms, Meertens Instituut, KNAW Amsterdam*.
- Paster, Mary Elizabeth (2006). *Phonological Conditions on Affixation*. Ph.D. thesis, University of California.
- Stump, Gregory T. (2001). *Inflectional Morphology: A theory of Paradigm Structure*. Cambridge: CUP.
- Walker, Rachel & Bella Feng (2004). A ternary Model of Morphology-Phonology Correspondence. *Proceedings of WCCFL*, vol. 23, 773–786.