Catalan vowel epenthesis as evidence for the free ride approach to morphophonemic learning

Maria-Rosa Lloret
Clàudia Pons-Moll

Available at: https://works.bepress.com/claudia_pons-moll/24/
1 Introduction

Although the discovery of underlying representations (URs) is a complex task that may be conditioned by various factors (see, e.g., Albright 2002), it has traditionally been claimed to be quite straightforward when the UR is revealed by productive and transparent morphophonemic alternations and when there is a clear and regular phonological condition in the language that justifies the disparities between the presumed UR and the corresponding surface representation. Among other possible scenarios (see, e.g., Albright 2008, Kager 2008, Pater et al. 2012), uncertainty appears when these morphophonemic alternations do not exist, when they are not fully productive and transparent, or when more than one interpretation of the UR is possible. This is the case of the URs of words beginning with /HesC/H in Catalan, which we discuss here.

The purpose of this squib is twofold. On the empirical side, we provide new arguments for the epenthetic nature of the vowel in /HesC/H-words, arising from the interaction between potential word-initial vowel epenthesis (see section 2) and underapplication of vowel reduction (VR) in Majorcan Catalan (see section 3). On the theoretical side, we show how Majorcan Catalan learners take a “free ride” (McCarthy 2005b) in the process of constructing the UR of nonalternating forms involving /esC/ initials by generalizing the pattern—and the subsequent input-output mapping—observed in cases with transparent morphophonemic alternations and a similar syllabification problem (see section 4).

This research was supported by the Spanish Ministry of Economy and Competitiveness (FFI2013-46987-C3-1-P), the Catalan Government (2014SGR918), and the RecerCaixa 2011 program. Previous work on this topic was presented at the 10th Old World Conference in Phonology (2013) and the 22nd Manchester Phonology Meeting (2014). We thank the audiences at these events for discussion of various points. We have also benefited from very helpful comments by the anonymous LI reviewers, the Squibs and Discussion editors, John McCarthy, Marc van Oostendorp, Donca Steriade, Francesc Torres-Tamarit, and Max W. Wheeler. Errors are ours.
The Challenge and the Potential Solution

In Catalan, vowel epenthesis has often been invoked to explain the presence of the default vowel (e: [ə] in the Majorcan dialect under discussion) when its absence would entail the occurrence of a structure defying syllable well-formedness (see, e.g., Wheeler 1975, Mascaro 1976, Bonet and Lloret 1998, 2005, Lloret 2002). However, there is a contrast between the weakness of morphophonemic arguments that support word-edge epenthesis, both word-initially and word-finally (Wheeler 2005, Lloret and Jiménez 2007), and the more conclusive evidence for vowel epenthesis in word-internal position (Pons-Moll 2005, Wheeler 2005) and clitic groups (Bonet and Lloret 2005, Wheeler 2005). For our interests, we restrict the discussion to the contrast between initial and internal positions.1

In word-initial position, vowel epenthesis has been claimed to occur in cases like [əspərə] (1a), [əsθələ] (2a), and [əspət] (3a) as a strategy to avoid word-initial [sC]- clusters. But the θ ~ [ə] alternation that justifies the epenthetic nature of this vowel is only evident—and still only partially so—in the morphophonemic alternations found in a limited set of cases, as in [əspərə] ~ [prosperə] (1b), which are dubiously morphologically complex because their compositional meaning is not synchronically transparent and because recent (transparent) prefixed forms show the vowel [ə] after the prefix, as in [iŋsperət] (1c). In contrast with the data in (1), all other words, whether they are native (2) or loans (3), lack these alternations, and [ə] systematically appears in all corresponding prefixed forms, as in [superaθələ] (2b) and [superaθət] (3b).

1 a. esperar [əsperə] ‘to expect’
   b. proθesperar [prosperə] ‘to prosper’
   c. inθesperat [inθesperət] ‘unexpected’
   d. desθesperar [θəsperə] ‘to despair’

2 a. əscələ [əsθələ] ‘stairs’
    b. superθələ [superaθələ] ‘superstairs’
    c. hiperθələ [iperθələ] ‘hyperspecial’

3 a. espot [əspət] ‘spot’
    b. superespot [superaθət] ‘superspot’
    c. subsθələɡən [supθələɡən] ‘subslogan’

Only one small piece of empirical evidence supports the traditional claim for the epenthetic nature of this initial vowel. As Wheeler (2005)

---

1 The vowels under discussion (whether epenthetic or not) are underlined throughout the squib for the sake of clarity. The transcriptions correspond to the Majorcan Catalan varieties that display the palatal obstruents [c], [j], and [θ] in specific contexts, corresponding to the velars of the remaining varieties.
notes, replicating Harris’s (1979) argument for Spanish, the oxytone stress pattern of the verb *estar* ‘to stay’ in the forms of the 2nd and 3rd person singular and the 3rd person plural present indicative (i.e., 2sg [əstəs], 3sg [əstə], 3pl [əstən]) supports the epenthetic nature of the initial vowel, because the unstressed character of the inflectional suffixes of these forms regularly causes the stem to bear the stress (as in 3sg [əntra], from /əntra + a/ ‘(he) enters’), but does not do so in *estar* (e.g., 3sg [əstə], from /stə/). In his discussion of the need for constraints on URs, however, McCarthy (1981:244) remarks that “[t]he apparently irregular final stress, confined as it is to three forms, is not compelling, nor is the distributional gap of #e̞sC, since it is not without exception” (cf. Spanish [e̞stə] ‘this’).

Contrariwise, word-internal epenthesis to avoid an ill-formed syllable contact (Vennemann 1988) in certain verbal forms is claimed to be uncontroversial (e.g., [təms] in (4b)) and with respect to parallel inflected forms of verbs that belong to the same conjugation but whose URs do not give rise to potential syllabic problems (e.g., [romprə] in (4c)), and also because there exist regular phonological conditions that drive the insertion of the epenthetic vowel (Lloret 2002, Pons-Moll 2005, 2011a, Wheeler 2005, Lloret and Jiménez 2007).

(4) a. teme-riə /teme-re / [təmɑɾe] / [təməɾiə] ‘(I) will/would be afraid’
creixe-riə /creixe-re / [kɾaʃəɾe] / [kɾaʃəɾiə] ‘(I) will/would grow up’
b. tem-θ /tem-θ-ies / [təms] / [təms] ‘(you) are/were afraid’
creixθ /creixθ-ia / [kɾeʃ] / [kɾaʃiə] ‘(he) grows/grew up’
c. rompθ-riə /rompθ-ria / [ɾomprə] / [ɾomprə] ‘(I) will/would break’
beuθ-riə /beuθ-ria / [bəwɾe] / [bəwɾiə] ‘(I) will/would drink’

The key question that we address here is whether the learner makes use of these word-internal alternations (and the subsequent fully legitimized unfaithful /θ/ → [a] map) to project the same mapping in the word-initial cases presented in (1)–(3). As we show next, the patterns related to the underapplication of VR found in words beginning with əsC- in Majorcan Catalan demonstrate that this is the case.

3 Underapplication of Vowel Reduction in Majorcan Catalan

Most Majorcan Catalan varieties have eight vowels in stressed position (i, e, ə, a, ɔ, ʊ) and four vowels in unstressed position (i, ə, ɔ, u). This system results from a general process of VR, according to which in unstressed position the mid-front vowels (e, ə) and the open-central vowel (/a/) merge as [ə], while the open-mid back vowel (/ɔ/) becomes close-mid ([o]) (see, e.g., Bibiloni 1983, Mascaro 2002, Wheeler 2005). The reduction to schwa explains the vocalic alternations illustrated in (5a–b), as well as the absence of [e], [ɛ], and [a] in unstressed position in the cases without alternations (6).
(5) a. **Stressed position**

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>casa</td>
<td>[cəˈzə]</td>
<td>‘house’</td>
</tr>
<tr>
<td>cafè</td>
<td>[caˈfɛ]</td>
<td>‘coffee’</td>
</tr>
<tr>
<td>carrer</td>
<td>[kəˈrɛr]</td>
<td>‘street’</td>
</tr>
<tr>
<td>contest</td>
<td>[kənˈtɛst]</td>
<td>‘(I) answer’</td>
</tr>
<tr>
<td>pedac</td>
<td>[pəˈdɑs]</td>
<td>‘dishtowel’</td>
</tr>
</tbody>
</table>

b. **Unstressed position**

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>casa</td>
<td>[caˈzətə]</td>
<td>‘house.DIM(INUTIVE)’</td>
</tr>
<tr>
<td>cafet</td>
<td>[caˈfetət]</td>
<td>‘coffee.DIM’</td>
</tr>
<tr>
<td>carreró</td>
<td>[kəˈrɛɾo]</td>
<td>‘street.DIM’</td>
</tr>
<tr>
<td>contestam</td>
<td>[kənˈtɛstəm]</td>
<td>‘(we) answer’</td>
</tr>
<tr>
<td>vellut</td>
<td>[veˈlut]</td>
<td>‘velvet’</td>
</tr>
</tbody>
</table>

VR of [ɛ] and [e], though, fails to apply under three circumstances (see, e.g., Veny 1962, Bibiloni 1998, Mascaro 2002, 2005, Wheeler 2005): (a) in productive derived forms with an unstressed vowel that alternates with a stressed mid-front vowel ([ɛ], [ɛ]) in the stem of the base (7a); (b) in first conjugation verbal forms (by far the most productive conjugation) with an unstressed vowel that alternates with a stressed close-mid front vowel ([ɛ]) in other verbal forms of the same inflectional paradigm (7b); and (c) in loans and learned (erudite) words with an unstressed mid-front vowel, generally preceded by a labial consonant (7c).

Pons-Moll (2011b, 2012, 2013) shows that the three contexts in which underapplication of VR occurs involve both paradigmatic pressure and morphological productivity of the related words, (7a–b), or just the intrinsic productivity that characterizes loans and learned words, (7c). Moreover, all the cases crucially have in common that the unreduced vowel is always—and only—located in the initial syllable of the stem.

(7) a. **Underapplication in productive derivation (native lexicon)**

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>festassa</td>
<td>[fesˈtasa]</td>
<td>‘party.AUGMENTATIVE’</td>
</tr>
<tr>
<td>ceglet</td>
<td>[səˈlɔt]</td>
<td>‘sky.DIM’</td>
</tr>
</tbody>
</table>

b. **Underapplication in verbal inflection (native lexicon)**

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>pegau</td>
<td>[peˈjaw]</td>
<td>‘to hit’</td>
</tr>
<tr>
<td>quedam</td>
<td>[kəˈdəm]</td>
<td>‘we stay’</td>
</tr>
</tbody>
</table>

This positional requirement is corroborated by two decisive patterns in Majorcan Catalan. First, when unstressed (e) is not located in the initial syllable of the stem, VR straightforwardly applies in productive derived words (e.g., [kərɛro] in (5b) vs. [fesˈtasa] in (7a)), in verbal forms (e.g., [kənˈtɛstəm] in (5b) vs. [peˈjaw] in (7b)), and in loans and learned words (e.g., amenitzar [aˈmanidzəɾ] ‘to liven up’, prefegrent [preˈfəɾənɾ] ‘preferable’ vs. [veˈdɔt] in (7c)). Second, in prefixed words, VR does not apply in the second unstressed (e) but does in the first one, as in renegat [reˈnetət] ‘great-grandson.DIM’ (cf. netet [neˈtet] ‘grandson .DIM’), recregmar [rekɾəˈməɾ] ‘to burn again’ (cf. crecrement [kɾəˈməɾ] ‘to burn’), afeminat [afəˈminat] ‘effeminate’ (cf. feminisme [fəˈminisme] ‘feminism’).

In Pons-Moll 2011b, 2013, the cases of underapplication of VR in productive derivation (7a) and verbal inflection (7b) are formally
accounted for through a set of output-output faithfulness constraints—the former cast within the Transderivational Correspondence Theory (Benua 1997) and the latter within the Optimal Paradigms model (McCarthy 2005a)—that target the vowel of the initial syllable of the stem (for present purposes, we will ignore the productivity effect). Underapplication in loans and learned words (7c), for which there are no alternations and thus no bases to exert pressure, is formalized in Pons-Moll 2012 through a contextual markedness constraint that penalizes a schwa in stem-initial position and is active only in the productive phonology (this limitation is necessary to capture the contrast between the forms in (7c), where VR underapplies, and the forms in (6), where VR applies regularly). 2

4 The Interaction between Word-Initial Epenthesis and Underapplication of Vowel Reduction

4.1 The Facts

Of crucial interest for our argument is the fact that underapplication of VR in the second (e) is also found in productive derived forms and verbal forms with (asCe)- ([asCe]-) initials, where the second (e) alternates with a stressed vowel ([é]), (8), and in loans and learned words beginning with (esCe)- ([esCe]-), (9).

    esquemet [esˈkemət] ‘scheme’ cf. esquema [esˈkema] ‘scheme’
    esperau [esˈperaʊ] ‘(you) wait’ cf. espera [əˈspéra] ‘(he) waits’

(9) especial [esˈpesjal] ‘special’ spermatozou [ˈspɜrmətɔzoʊ] ‘spermatozoan’
    especific [esˈpesifik] ‘specific’ estgreotip [ˈesɡreotip] ‘stereotype’

4.2 Empirical Consequences

Since the generalization for underapplication of VR is that it only occurs when the vowel is located in the initial syllable of the stem (see section 3), we now have external and independent evidence that lies inside the grammar to assert that the initial vowel of words beginning with [asC]- is not part of the stem and should therefore be considered epenthetic: if this vowel were part of the stem, the second (e) would be located in a noninitial syllable of the stem and hence would not be affected by VR underapplication (but would instead reduce, as it does in the aforementioned non-initial-syllable cases). The initial

---

2 Note, incidentally, that, since the output-output faithfulness constraints and the contextual markedness constraint posited in this analysis are sensitive to morphological edges, the proposal hinges on a “containment” approach to faithfulness within Optimality Theory (OT).
vowel of [asC]- words behaves, in fact, as if it were “invisible” to the output-output positional faithfulness constraints alleged to account for VR underapplication in derivation (cf. [festása] ~ [fěstə], and also [astevát] ~ [astévo]) and verbal inflection (cf. [pejáw] ~ [pejja], and also [asperáw] ~ [aspēra]), and it is unaffected by the contextual markedness constraint against a schwa in the initial syllable of the stem presumed to account for VR underapplication in loans and learned words (e.g., [bënina], [feitufízma], [vegőt], and also [aspesjal]). All in all, this is clear positive evidence for the epenthetic nature of the initial vowel: if it were not epenthetic, the second (e) would not be targeted and affected by these constraints, because it would occur in a position other than the initial one within the stem.

4.3 Theoretical Consequences

The data analyzed are also relevant to test three of the approaches developed within OT to account for the nature of URs and their process of acquisition, learning, and construction: richness of the base (ROTB) and lexicon optimization (LO) (Prince and Smolensky 2004), and the free ride approach to morphophonemic learning (FRML) (McCarthy 2005b).3

Under the ROTB hypothesis and guided only by morphophonemic alternations, the potential URs for words like [astevát], [asperáw], [aspesjal] (without the alternation [0]- ~ [σ]- and with the second vowel unreduced) should be /0sC/- and /asC/-, Clearly, though, if we assume a UR with the vowel (i.e., /asC/-), we would obtain nonexistent forms with reduction of the second (e) to schwa (e.g., *[astevóv], *[asperáw], *[aspesjal]). This is so because the output-output faithfulness constraints that target the vowel of the initial syllable of the stem (adduced to explain cases like [festása], [pejáw]) or the contextual markedness constraint against a schwa in the initial syllable of the stem (adduced to explain cases such as [feitufízma]) would not affect the second vowel of words beginning with [asC]-, as it would not be located in the initial syllable of the stem: *[astevát], *[asperáw], *[aspesjal]. This does not counter the ROTB hypothesis in a general way, but limits its application to the data we are dealing with here. Under the LO hypothesis, words like [astevát], [asperáw], and [aspesjal] are expected to have URs of the shape /asC/-, Hence, the same reasons provided for the ROTB hypothesis invalidate this approach.4

3 A thorough survey of learning models in phonology, with specific reference to phonological alternations, can be found in Albright and Hayes 2011, and a detailed discussion about the free ride proposal and its consequences can be found in Krämer 2012.

4 See, additionally, Nevins and Vaux 2008 for a detailed discussion of the flaws of LO, based on empirical and experimental grounds, and Vaux 2005 for the theoretical and empirical problems related to ROTB and LO.
So it seems that learners take a free ride with the unfaithful /θ/ → [a] map. The FRML presumes that “[w]hen alternation data tell the learner that some surface [B]s are derived from underlying /A/s, the learner will under certain conditions generalize by deriving all [B]s, even non-alternating ones, from /A/s,” and “take a ‘free ride’ on the /A/ → [B] unfaithful map” (McCarthy 2005b:19). Thus, in the absence of (fully productive) morphophonemic [θ] → [a] alternations that shed light on the URs of words with [asC]- initials (and also given the fact that learners have limited experience and are often unaware of the morphophonemic alternations that would allow them to discover URs), speakers generalize an unfaithful /θ/ → [a] map, which they deduce from the cases in which there are dynamic alternations (e.g., [tέms], [romprέ] ~ [tόmarέ]), and extend it to nonalternating cases (e.g., [stevέt], [asperάw], [aspesjάl]).

Significantly, our proposal matches McCarthy’s (2005b) proposal according to which learners take the free ride strategy in nonalternating forms under certain conditions only: namely, when, by generalizing the unfaithful map, a “consistent” and “more restrictive” grammar (p. 21) than the one obtained by an identity map (as in LO) is achieved. This is exactly the case we are dealing with here. First, the grammar obtained by generalizing the unfaithful /θ/ → [a] map to all cases, including nonalternating ones, is consistent with the primary data, because it homogenous explains the complete absence of words beginning with [sC]- in the language (both those with alternations and those without) via the constraint *sC-, and it further limits the number of input-output maps. Second, the grammar obtained is more restrictive than the one obtained by an identity map because it has a higher r-measure. According to Prince and Tesar (2004:252), “[t]he r-measure for a constraint hierarchy is determined by adding, for each faithfulness constraint in the hierarchy, the number of markedness constraints that dominate that faithfulness constraint.” A grammar that grants “more power to markedness constraints” is, therefore, “more restrictive” (McCarthy 2005b:32).

In the proposed grammar for Majorcan Catalan, the identity /æ/ → [a] map and the subsequent grammar of the earliest stages (see (10)) are abandoned, because the morphophonemic [θ] → [a] alternations are discovered and incorporated into the grammar (see (12)). And once the free ride has taken place and the ensuing unfaithful map /θ/ → [a] has been generalized (see (11)), not only the markedness syllabic constraints SyllableContact (Vennemann 1988), which penalizes syllabic contacts like *[tόmrέ], and MinimumSonorityDistance (Steriade 1982, Clements 1990), which avoids onset sequences with low sonority distances such as *[tόmrέ], must be ranked above the faithfulness constraint against insertion, DEP₁₀ (see (12)), but *sC- must be ranked above it as well, to justify vowel insertion in these cases also (see (11)). The promotion of *sC- above DEP₁₀, along with SyllableContact and MinimumSonorityDistance, indeed implies a grammar that gives more power to markedness constraints and hence has a higher r-measure; compare (10) with (11).
These tableaux show how the free ride mechanism works. The old support tableau in (10) reflects a grammar and an input-output map based exclusively on the phonotactics of the language (phonotactic learning), where there is no room for morphophonemic alternations. The new support tableau in (11) represents a grammar in which morphophonemic alternations have already been discovered (morphophonemic learning), leading to the construction of the unfaithful map /θ/ → [a] and the following “surgery” of the previously constructed UR /asC/-. By virtue of the free ride, the surgery gives rise to the new UR /sC/-.

Summing up, once the identity /a/ → [a] map of the earliest stages is abandoned because the morphophonemic [θ] ~ [a] alternations are discovered and incorporated into the grammar, the learner is fully committed to the unfaithful /θ/ → [a] map, which applies to all cases with a parallel syllabic problem (thus following the orthodox, not contextually determined, across-the-board free ride) and which is driven by the constraint ranking [SYLLABLECONTACT, MINIMUMSONORITYDISTANCE, *SC- >> DEPIO], with all syllabic well-formedness markedness constraints outranking faithfulness. Looking forward, these data point to
the conclusion that unfaithful maps may be legitimized not only by their own morphophonemic alternations but also by the neighboring ones, provided that a more restrictive grammar is achieved.

References


Pons-Moll, Clàudia. 2012. Loanword phonology, lexical exceptions, morphologically driven underapplication, and the nature of po-


