What can children's spelling tell us about underlying representations?[*]

Clàudia Pons-Moll[1] and Gisela Fuertes[2]

[1]*University of Barcelona,* [2] *INS El Palau (Departament d'Educació)*

1   Introduction

This paper deals with a perennial debate in phonology, focusing specifically on the phonology of Catalan. What is the underlying representation of non-alternating segments? In the case of Catalan, what is the underlying representation (UR) of the surface [ə]s and [u]s that never alternate with a vowel in stressed position? Which theoretical approach provides a more realistic answer to these questions?

In order to shed some light on this empirical and theoretical debate, in this article we present the results of an experimental survey analyzing the spelling of first and second grade primary school children. First grade children are at a very early stage of the process of learning to read; they have had thus limited exposure to conventional spelling and their knowledge of it is unstable. In contrast, second grade primary school children have had more exposure to conventional spelling (Varnhagen et al. 1997) and are more aware of the morphological relations among words. The specific goals of the paper are to identify how first and second graders spell non-alternating and alternating [ə] and [u] in Central Catalan, and to establish whether there are differences in the way they spell these vowels. We worked with three factors: *a)* the type of vowel involved (i.e. [ə] or [u]); *b)* the alternating or non-alternating character of the involved vowels; *c)* the age of the children.

Although orthography has traditionally been considered to have nothing to do with phonological theory or phonological knowledge, we hypothesized that the way primary school children spell unstressed vowels may help to identify the UR to which they associate these vowels. In fact, some previous studies (see, among others, Treiman et al. 1994, Treiman 2017) have indeed proved that phonology, morphology, and orthography are closely related systems that interact during development. Treiman (2017: 6), for instance, collects robust evidence on the influence of phonology on children's spelling: "As an example of how several types of knowledge can influence young children's letter choices, consider the taps that occur in the American English pronunciations of words such as *city*, *spider*, *eater*, and *reading*. A tap is a quick flap of the tongue just behind the upper teeth. It is voiced, meaning that the vocal cords vibrate during its production. In this respect, taps are more similar to /d/, which is voiced, than to /t/, which is not voiced. When offered a choice between ‹d› and ‹t› to spell a tap, U.S. 5- and 6-year-olds showed an overall preference for ‹d› (Bourassa, Treiman, & Kessler, 2006; Treiman, Cassar, & Zukowski, 1994; Wolter, Wood, & D'zatko, 2009). Taps are more likely to be spelled as ‹t› than as ‹d› in English words, so children's preference for ‹d› points to an influence of phonology.". As for the influence of morphology on spelling, the same author claims: "Children in these studies were also influenced by the morphological structure of the word [...]. They were more likely to choose ‹t› for words like *eater*, which have a stem that ends with /t/, than for words like *city*, which do not have such a stem. Use of ‹t› was least common for words like *reading*, where the voicing of the tap (phonology) encourages uses of ‹d› and where the existence of a stem with a final /d/ spelled as ‹d› (morphology) also encourages this choice." (Treiman 2017: 6). In a different vein, Hamann & Colombo (2017) have shown that orthography, perception and phonology also interact in the process of loanword adaptation from English to Italian.

That children use phonological and morphological knowledge in their spelling choices is an hypothesis we worked with in this study, and, as we will show, it is generally confirmed by the results.

The paper is organized as follows. In § 2 we present the data relative to alternating and non-alternating vowels which have aroused some controversy in Catalan phonological studies. In § 3 we focus on the theoretical debate, framed within Optimality Theory (OT), concerning the determination of URs when no dynamic alternations are available to the learners. In § 4 we sketch out the goals and the related hypotheses of the study. In § 5 we describe the survey we conducted, as well as the methodology we followed. In § 6 we present the results and the main findings of the survey, and in § 7 we conclude.

2   Empirical debate

In Eastern Catalan, surface unstressed [ə]s can alternate with a stressed [á], [ɛ́] or [é] (1a), and unstressed [u]s can alternate with a stressed [ú], [ɔ́] or [ó] (1b). This has traditionally been taken as evidence of an UR of these occurrences with /a/, /ɛ/, /e/, /u/, /ɔ/, and /o/ respectively. In some cases, though, unstressed [ə]s and [u]s do not have the chance to alternate with a vowel in stressed position, because the stem in which they appear is always unstressed (2a; 2b), and so the UR remains uncertain. Unstressed [ə]s may correspond to an underlying /a/, /ɛ/, /e/ or /ə/, and unstressed [u]s may correspond to an underlying /ɔ/, /o/ or /u/.

(1) Vowel alternations for [ə] and [u] in Eastern Catalan

    a.   Alternating cases for [ə]   certain URs

        c[ə]*seta* 'house dim.'~ *c*[á]*sa* 'house' → /a/

        p[ə]*lut* 'hairy'~ *p*[ɛ́]*l* 'hair' → /ɛ/

        d[ə]*nteta* 'tooth dim.'~ *d*[é]*nt* 'tooth' → /e/

    b.   Alternating cases for [u]   certain URs

        ll[u]*cet* 'hake dim.'~ *ll*[u]*ç* 'hake' → /u/

        b[u]*leta* 'ball dim.'~ *b*[ɔ́]*la* 'ball' → /ɔ/

        p[u]*meta* 'apple dim.'~ *p*[ó]*ma* 'apple' → /o/

(2) Lack of vowel alternations for [ə] and [u] in Eastern Catalan

    a.   Non-alternating cases   uncertain URs

        m[ə]*rtell*   'hammer' → /a/, /ɛ/, /e/, /ə/?

        r[ə]*spall*   'brush' → /a/, /ɛ/, /e/, /ə/?

    b.   m[u]ssol   'owl' → /u/, /ɔ/, /o/?

        b[u]txaca   'pocket' → /u/, /ɔ/, /o/?

Between the 50s and the 80s, the UR of non-alternating [u]s and especially [ə]s was the object of an endless debate among scholars working on Catalan phonology, which proved to be unproductive in many respects (see Mascaró 1991). Within structuralism, Alarcos (1953, 1973), for instance, proposed the archiphonemes /A/ and /U/ for non-alternating [ə] and [u] (cf. *m*/A/*rtell* 'hammer'; *m*/U/*ssol* 'owl'), whereas Badia (1951, 1965) defended the phonemic character of /ə/ in monosyllabic words. From the perspective of generative phonology, Mascaró (1978) and Wheeler (1979) arbitrarily assumed an underlying /a/ for non-alternating [ə] (cf. *m*/a/*rtell* 'hammer'), while Viaplana & DeCesaris (1984) argued for an underlying /ə/. Bonet & Lloret (1998) gave arguments for either /ə/ and /u/ and for /A/ and /U/ as URs for non-alternating [ə] and [u]. Overall, as the arguments for each position were generally weak, or, at least, difficult to prove, the UR of non-alternating vowels remained difficult to untangle, and the whole controversy became rather sterile (see Mascaró 1991).

3   Theoretical debate

OT has developed various (competing) theories and hypotheses about the nature of URs and the process of their acquisition, construction, storage, and access when no dynamic morphophonemic alternations are (yet) available to the speaker-learner or to the analyst. These theories include Richness of the Base (Prince & Smolensky 2004), Lexicon Optimization (Prince & Smolensky 2004; Smolensky 1996) and the Free-Ride in Morphophonemic Learning (McCarthy 2005). The last of these, as we will see, is challenged when the mapping between a surface representation and its corresponding UR is not univocal, as in the case of the examples presented in (2a,b).

According to the Richness of the Base hypothesis (RoTB) (Prince & Smolensky 2004: 209), in cases lacking alternations, the analyst must project all possible URs for every surface form (Prince & Smolensky 2004: 209). The grammar, that is, the constraint hierarchy, is ultimately responsible for selecting the actual surface form in a given language, no matter which UR is taken.

According to Prince & Smolensky (2004), however, in the process of storage and access to URs, the Lexicon Optimization principle is assumed to be at play (Prince & Smolensky 2004: 209-210). This principle establishes that when there is no morphophonemic evidence bearing on the choice of URs, phonological representations are stored identically to their surface form, leading to a direct economization of input-output mappings (given that the map from underlying to surface representations is accomplished more faithfully).

In relation to these hypotheses, it is worth mentioning the criticism made by Bermúdez-Otero (2006): "[…] in strictly parallel versions of OT, once the phonologist has satisfied himself (i) that the constraint hierarchy generates wellformed outputs for every possible input and (ii) that there is a viable input for every output, he has little incentive to ask what input representation is actually selected by the learner and how, crucial though these questions are to the psycholinguist and to the historical linguist." (Bermúdez-Otero 2006: 503). Due to the output-oriented character of OT, indeed, the exact UR of non-alternating segments is of little interest to the phonologist, provided that the constraint hierarchy leads to the selection of the desired, actual, candidate.

In some particular cases, though, it is possible to find external evidence for a specific UR, despite the lack of direct morphophonemic alternations. This is what is assumed by McCarthy (2005)'s proposal of the free-ride in morphophonemic learning. According to this theory, when alternation data tell the learner that some surface [B]s are derived from underlying /A/s, s/he will under certain conditions generalize by deriving all [B]s, even non-alternating ones, from /A/s, so that "an adequate learning theory must […] incorporate a procedure that allows non-alternating [B]s to take a "free ride" on the /A/ → [B] unfaithful map." (McCarthy 2005: 19). The conditions under which learners take the free ride strategy in non-alternating forms are the following: when, by generalizing the unfaithful map, a *a*) "consistent" and *b*) "more restrictive" grammar than the one obtained by an identity map is achieved (McCarthy 2005: 21). Following Prince & Tesar (2004: 252), "[t]he r[estrictiveness]-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", so that a grammar that grants "more power to markedness constraints" is "more restrictive" McCarthy 2005: 32).
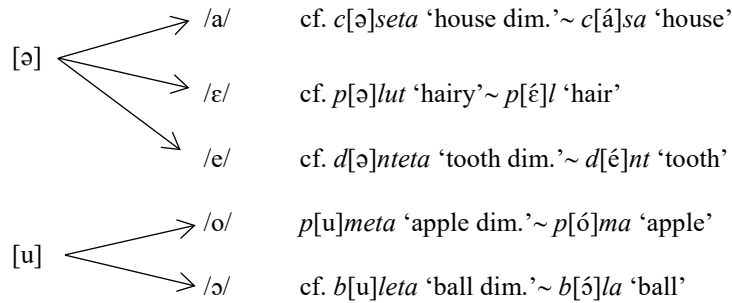
The free ride strategy has proven to be true, for instance, for cases of coalescence in Sanskrit, Choctaw and Rotuman (McCarthy 2005), for cases of hyperrhoticity in some varieties of English (Krämer 2012), or for vowel epenthesis in Majorcan Catalan, for which there is independent evidence, based on its interaction with the underapplication of vowel reduction, that learners generalize the unfaithful map /∅/ → [ə], derived from dynamic morphophonemic alternations, to non-alternating items (Pons-Moll & Lloret 2014, Lloret & Pons-Moll 2016). However, as argued for in Pons-Moll (2016), the free ride strategy is challenged when the input-output mapping(s) derived from dynamic alternations and which are potentially generalized to non-alternating items are *not univocal*, that is, when the alternating [B]s derive from more than one UR, as in the cases of (2), which we reproduce in (3).

(3) The free ride strategy challenged (based on Pons-Moll 2016)

    a.   Non-alternating [Bs]

m[ə]*rtell*   'hammer'
m[u]*ssol*   'owl'

b.   Alternating [Bs]…   derived from /A/, /C/, /D/

[ə] ⟶ /a/   cf. *c*[ə]*seta* 'house dim.' ~ *c*[á]*sa* 'house'

[ə] ⟶ /ɛ/   cf. *p*[ə]*lut* 'hairy' ~ *p*[ɛ́]*l* 'hair'

 ⟶ /e/   cf. *d*[ə]*nteta* 'tooth dim.' ~ *d*[é]*nt* 'tooth'

[u] ⟶ /o/   *p*[u]*meta* 'apple dim.' ~ *p*[ó]*ma* 'apple'

[u] ⟶ /ɔ/   cf. *b*[u]*leta* 'ball dim.' ~ *b*[ɔ́]*la* 'ball'

As we can see in these examples, morphophonemic alternations do not provide any clues as to the nature of the UR, because the relation between the surface representation and the UR is not one-to-one but one-to-many. So the free-ride strategy does not seem possible in cases of this type (see Pons-Moll 2016, however, for a possible theoretical solution for consumating the "free ride" in these types of situations, based on efficiency in terms of violations of ranked faithfulness constraints).

4   Goals and hypotheses

The main goal of this study is to establish how first grade and second grade primary school children (aged, respectively, between 6 and 7 and 7 and 8) spell non-alternating and alternating [ə] and [u] of Central Catalan. This goal is based on the hypothesis that the way in which first graders spell unstressed vowels may help to identify the UR to which they associate these vowels, because they are at a very early stage of the process of learning to read, and are unfamiliar with conventional spelling; they are expected to use, thus, another kind of knowledge, including phonological knowledge. A second goal of the study is to determine whether there are differences in the spelling of alternating and non-alternating vowels. If no significant differences are found, this might mean that morphophonemic alternations are not yet considered at these stages of phonological acquisition, thus lending support to the Lexicon Optimization hypothesis. If, in contrast, there are significant differences, this might mean that morphophonemic alternations are already considered at these stages. Finally, a third goal of this study is to identify any possible differences in the spelling between first and second graders in this respect. If significant differences are found, this might demonstrate the stage at which morphophonemic alternations, along with conventional spelling, are starting to be taken into account.

Note that in order to avoid the obvious bias of conventional orthography in the spellings the children provided, we were guided basically by the misspellings rather than by the correct spellings. A correct spelling can reflect a certain pronunciation, a certain UR beyond this pronunciation, but also an awareness of the conventional spelling. An incorrect spelling for a given word, on the other hand, conveys that the orthographic conventions have not yet been assimilated, and so other factors are at play: among them, the pronunciation of the word and the abstract representation the children have constructed for this word.

Overall, we envisage that the way children spell non-alternating vowels might shed some light on the nature of URs and the process of their acquisition, construction, storage, and access when no dynamic morphophonemic alternations are (yet) available to the learner.

5   Survey

A group of 29 primary school children aged between 6 and 7 (attending first grade) and a group of 39 primary school children aged between 7 and 8 (attending second grade), from the Escola Sant Martí and the Escola Can Coll and with (Eastern) Catalan as L1, were asked to answer a test in which they had to write down the words corresponding to 21 images projected on a large screen using PowerPoint. The selected

words were considered to be appropriate for the children's stage of development, and presented the same degree of difficulty in terms of meaning and morphological complexity.

Ten of the elicited words included orthographic *a* and *e* corresponding to unstressed [ə]s alternating with either [á], [ɛ́] or [é] (4a) and orthographic *u* and *o* corresponding to unstressed [u]s alternating with either /o/ and /ɔ/ (4b).

(4)

    a.   Selected words with alternating [ə]

Spelling	Phonetics	Alternation
caseta 'house dim.'	*c*[ə]*seta*	cf. *c*[á]*sa* 'house'
palet 'stick dim.'	*p*[ə]*let*	cf. *p*[á]*l* 'stick'
pelut 'hairy'	*p*[ə]*lut*	cf. *p*[ɛ́]*l* 'hair'
pereta 'pear dim.'	*p*[ə]*reta*	cf. *p*[ɛ́]*ra* 'pear'
denteta 'tooth dim.'	*d*[ə]*nteta*	cf. *d*[é]*nt* 'tooth'
peixet 'fish dim.'	*p*[ə]*ixet*	cf. *p*[é]*ix* 'fish'

    b.   Selected words with alternating [u]

Spelling	Phonetics	Alternation
osset 'bear dim.'	[u]*sset*	cf. [ó]*s* 'bear'
pometa 'apple dim.'	*p*[u]*meta*	cf. *p*[ó]*ma* 'apple'
osset 'bone dim.'	[u]*sset*	cf. [ɔ́]*s* 'bone'
boleta 'ball dim.'	*b*[u]*leta*	cf. *b*[ɔ́]*la* 'ball'

The other eleven words included orthographic *a* and *e* corresponding to non-alternating (at least transparently for the children) unstressed [ə]s (5a), and *o* and *u* corresponding to non-alternating (at least transparently for the children) unstressed [u]s (5b). We ended up excluding from the analysis one of the words considered in this group, **joguina* ('toy'), because, in fact, it does alternate with *joc* ('toy'), so the number of words with non-alternating vowels was ten.

(5)

    a.   Selected words with non-alternating [ə]

Spelling	Phonetics	Alternation
martell 'hammer'	*m*[ə]*rtell*	-----
raspall 'brush'	*r*[ə]*spall*	-----
lleó 'lion'	*ll*[ə]*ó*	-----
tresor 'treasure'	*tr*[ə]*sor*	-----
vestit 'dress'	*v*[ə]*stit*	-----

    b.   Selected words with non-alternating [u]

Spelling	Phonetics	Alternation
conill 'rabbit'	*c*[u]*nill*	-----
formiga 'ant'	*f*[u]*rmiga*	-----
muntanya 'mountain'	*m*[u]*ntanya*	-----
mussol 'owl'	*m*[u]*ssol*	-----
butxaca 'pocket'	*b*[u]*txaca*	-----

Let us now make some remarks on Catalan spelling, which are relevant to the interpretation of the data collected in the survey.

- The schwa ([ə]) is represented either by the graphemes *a* (as in *caseta* 'house dim.' and *palet* 'stick dim.') or *e* (as in *pelut* 'hairy', *pereta* 'pear dim.', and *denteta* 'tooth dim.').
- The high back vowel [u] is represented either by the graphemes *o* (as in *osset* 'bone dim.', *pometa* 'apple dim.', *osset* 'bear dim.', and *boleta* 'ball dim.') or *u* (as in *muntanya* 'mountain' and *mussol* 'owl').
- There is no specific grapheme for [ə]. This means that *a* can correspond either to [a] or to [ə].
- Except for some cases with graphic accentuation, there is no spelling distinction for [ó] and [ɔ] or for [ɛ] and [é], which correspond, respectively, to *o* and *e*.
- Most cases of alternation are reflected in the spelling, as in *caseta* 'house dim.', derived from *casa caseta* 'house', and *pelut* 'hairy', derived from *pèl* 'hair'.
- Cases without alternation generally take the spelling of the pronunciation of the word in non-reducing varieties, that is, in Western Catalan: *martell* 'hammer', *raspall* 'brush' are pronounced with [a] so they are spelled with *a*; *lleó* 'lion', *tresor* 'treasure', *vestit* 'dress' are pronounced with [e] so they are spelled with *e*; *formiga* 'ant' is pronounced with [o], so it is spelled with *o*; *muntanya* 'mountain', *mussol* 'owl', *butxaca* 'pocket' are pronounced with [u], so they are spelled with *u*.

As stated above, 21 images were projected on a large screen, and the children had to write down the corresponding words on a piece of paper, one after the other (see some samples of the answers in § 7). We limited the items to 21 in order to avoid distraction. The items with vowel alternations (4a,b) were elicited by projecting the large and small versions of each element (i.e., a drawing of a large house and a small version of the same house). The aim of this methodology was to prompt the children to relate the diminutives with their primitives. The items without vowel alternations (5a,b) were elicited by projecting the corresponding image (i.e., a drawing of a hammer, of a rabbit, etc.). The experiment was guided, but, of course, no cues for the spelling of the words were given. The answers for each item were transcribed on a grid, corresponding to each informant. The answers were classified in five categories: *1*) spelling with *a* for [ə] (in blue in the graphs below); *2*) spelling with *u* for [u] (also in blue in the graphs below); *3*) spelling with *e* for [ə] (in yellow in the graphs below); *4*) spelling with *o* for [u] (also in yellow in the graphs below); *5*) others (in gray in the graphs below). Afterwards, we calculated the percentages of answer for each category.

6   Results and discussion

In this section we present and discuss the results of the survey. In all the figures, blue indicates that the children wrote *a* for alternating and non-alternating unstressed [ə]s and *u* for alternating and non-alternating unstressed [u]s, and yellow that they either wrote *e* for alternating and non-alternating unstressed [ə]s or *o* for alternating and non-alternating unstressed [u]s. Finally, gray indicates other solutions provided by the children, such as the omission of the vowel, as in *dnteta* for *denteta* or in *pstis* for *pastís*; the occurrence of other vowels, possibly triggered by harmony processes, as in *pemeta* for *pometa*, in *llouo* for *lleó* or in *trosort* for *tresor*; the use of two vowels instead of one, as in *pealet* for *palet*; vowel metathesis, as in *trosar* for *tresor*; or a response with another lexical item, such as *branqueta* 'bough DIM.' instead of *palet*, *cofra* 'coffer' instead of *tresor*, *mostra* 'monster' instead of *pelut*, etc.

As indicated in § 4, in order to avoid the obvious bias or influence of conventional orthography in the spellings the children provided, when interpreting the results we were guided basically by the misspellings rather than by the correct spellings. A correct spelling can reflect a certain pronunciation, a certain UR beyond this pronunciation, but also the awareness of the conventional spelling. An incorrect spelling for a given word, on the other hand, conveys that the orthographic conventions have not yet been assimilated, and so other factors are at play: among them, the pronunciation of the word and the abstract representation the children have constructed for this word.

In Fig. 1 we present the results for first grade children (aged between 6 and 7). The most striking feature of these results is the lack of significant differences between the spelling of alternating and non-alternating vowels. Blue is the prevailing color for all items, meaning that the children generally spelled either alternating or non-alternating [ə] with *a* and either alternating and non-alternating [u] with *u*. If we

zoom in on the results, we can see, for instance, that the children chose *a* to spell the items *pelut* and *pereta* (alternating with [ɛ́]) and the items *lleó*, *tresor* and *vestit* (with non-alternating [ə] and spelled with *e*) in 81% and 82% of cases respectively. We also see that they chose *u* to spell the items *osset* and *boleta* (alternating with [ɔ́]) and the items *conill* and *formiga* (with non-alternating [u] and spelled with *o*) in 66% and 60% of cases respectively. These results suggest that morphophonemic alternations are not considered at this stage and that the kids chose the spelling which is closer to the pronunciation. More evidence for the fact that morphophonemic alternations are not considered is that they chose *o* to spell non-alternating [u] corresponding to *o* (*conill*, *formiga*) more frequently (33% of cases) than to spell [u] alternating with [ó] (*osset*, *pometa*) (21%); that is, presence of alternations does not encourage a correct spelling.

FIG. 1. *Results for first grade children (% of spelling solutions for each type of word considered)*
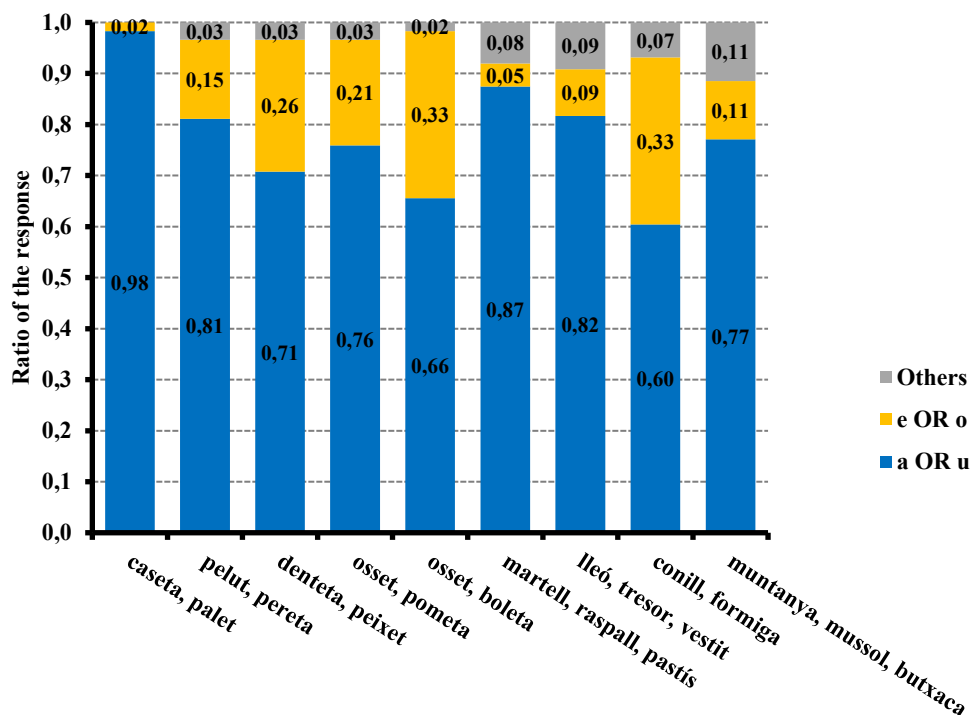
Among the non-alternating forms, the "better performance" with less misspellings for back vowels than for non-back vowels is striking: the items *conill* and *formiga* were correctly spelled, using *o*, in 33% of cases, whereas the items *lleó*, *tresor* and *vestit* were correctly spelled in only 9% of cases. The same asymmetry can be observed between the answers for *caseta* and *palet* and the answers for *muntanya*, *mussol* and *butxaca*: while the former are systematically spelled with *a*, the latter show more instances with *o*, although the conventional spelling is with *u*. This means that the children at this stage overgeneralize the mapping [u] → *o* more often than the mapping [ə] → *e*. What we infer from these results is that the alternations for back vowels are learned before the alternations for non-back vowels. The reason for this might be that the children have to deal with a more reduced typology of output-input alternations for back vowels (two to one) than for non-back vowels (three to one) (see 3b).

Among the words with alternating vowels, a "better performance" with less misspellings is achieved for unstressed [ə] alternating with the high-mid front vowel [é] (*denteta*, *peixet*) than for those alternating with the low-mid front vowel [ɛ́] (*pelut*, *pereta*). This might be because the dissimilarity between [e] and [ə] is greater than that between [ɛ́] and [ə], and is in agreement with the diachronic evolution of the stressed /ə/, which developed into the low-mid vowel [ɛ́] in many dialects of Catalan. Note also the slightly better performance for unstressed [u] alternating with the low-mid vowel [ɔ́] (*osset*, *boleta*), wrongly spelled with *u* in 76% of the cases, compared with those alternating with the high-mid vowel [ó] (*osset*, *pometa*), wrongly spelled with *u* in 66%. A plausible reason for this discrepancy is the greater dissimilarity between [ɔ] and [u] than between [o] and [u].

Another plausible explanation for the "better performance" for unstressed [ə] alternating with the high-mid front vowel [é] than for those alternating with the low-mid front vowel [ɛ́] mentioned just above is the fact the alternations with the low-mid vowel [ɛ́] included a derived form which was not a diminutive (*pelut 'hairy'*), and was thus more difficult to relate to the primitive. This last circumstance can be observed in the following four figures, with the specific results for each word with [ə] alternating with [ɛ́]. The word *pelut*, which is a derived form of *pèl* 'hair', but not its diminutive, was wrongly spelled with *a* in 93% of cases; in contrast, the diminutives were wrongly spelled with *a* less often: *pereta* and *denteta* were spelled with *a* in 69% of the cases, and *peixet* in 72%.

FIG. 2. *Results for* pelut *'hairy'*
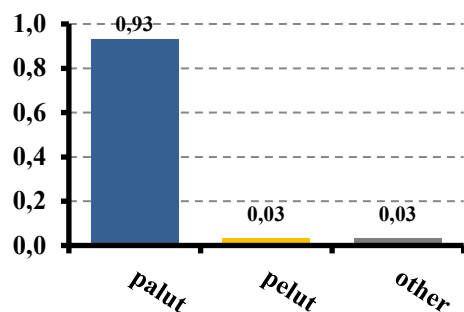
FIG. 3. *Results for* pereta *'pear* DIM.*'*
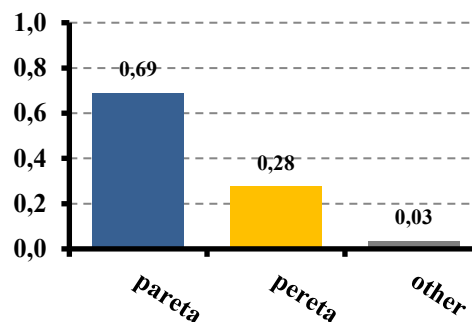
FIG. 4. *Results for* denteta *'tooth* DIM.*'*
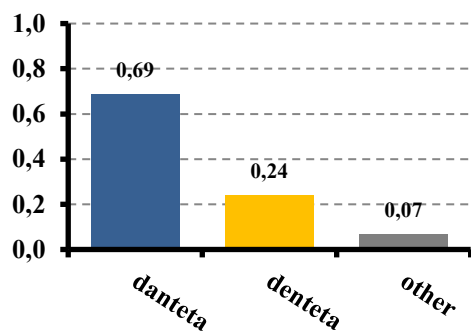
FIG. 5. *Results for* peixet *'fish* DIM.*'*
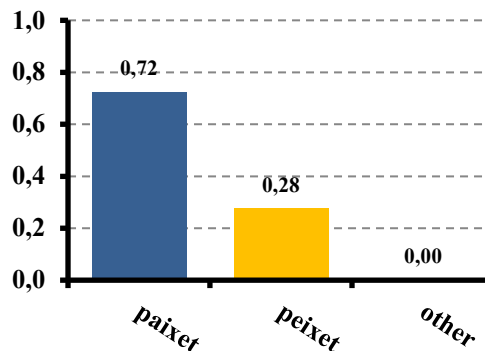
To conclude this section on the first graders' results, we stress that no significant differences were found among alternating and non-alternating [ə]s and [u]s: [ə] and [u] were predominantly associated with an orthographic *a* and *u* respectively, whether or not they alternate. These results generally support the Lexicon Optimization hypothesis, according to which there is a first stage in language acquisition in which, from all possible input candidates (Richness of the Base), the learner selects the one that matches the adult output representation as the optimal input (Smolensky 1996).

In Fig. 2 we present the results for second grade children (aged between 7 and 8). At this stage, the children show a clear improvement in spelling performance, especially for alternating vowels, which may mean that they take either conventional spelling or morphophonemic alternations into account. Yellow is the prevalent color for words with alternating mid vowels (*pelut*, *pereta*; *denteta*, *peixet*; *osset*, *pometa*; *osset*, *boleta*), and this means that they were correctly spelled, perhaps, as said, due to a consideration of the morphophonemic alternations. Another interesting observation is the fact that the results for second graders again show a better performance for back vowels than for non-back vowels, but in this case for alternating vowels. As we said above, this must be related to the fact that the schwa is involved in a more varied typology of alternations (one-to-three) than the unstressed [u] (one-to-two) (3b). The spelling mistakes for [ə] alternating with front mid vowels were again more prevalent in the vowels of the group *pelut*, *pereta* than in the vowels of the group *denteta*, *peixet*. Again we attribute this behavior to the presence of a non-diminutive (*pelut*) in the former group.

FIG. 6. *Results for second grade children (% of spelling solutions for each type of word considered)*
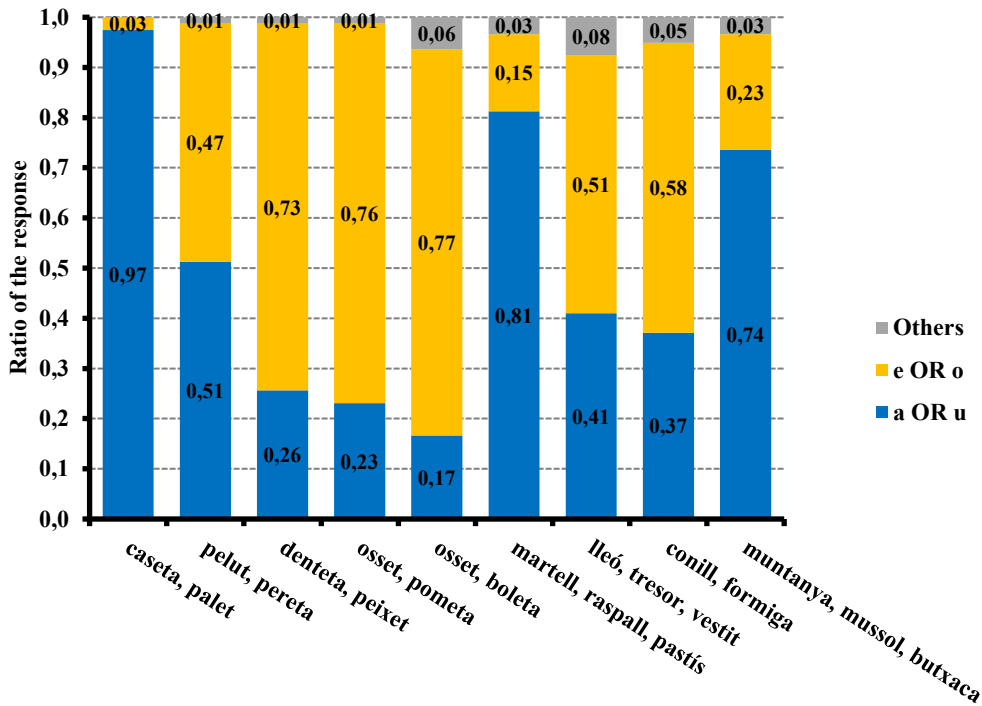
As reflected in the two figures below, however, the ability to spell unstressed schwas corresponding to *a* and unstressed [u] corresponding to *u* declines from first grade to second grade; that is, in first grade, these vowels were correctly spelled with *a* and with *u*, but in second grade, the same vowels start to be misspelled, with *e* and *o* respectively. In our view, this is plausible evidence that, once the morphophonemic alternations are incorporated (as demonstrated by the lack of mistakes for alternating [ə] and [u] in this grade), there is a stage of uncertainty or of recalculation of the UR corresponding to non-alternating vowels, probably supporting the free-ride approach to morphophonemic learning (McCarthy 2005). The reasoning is the following. Children at this stage have learned that some [ə] alternate with [ɛ́] and [é] and that some [u] alternate with [ɔ́] or [ó]; since in these cases the spelling is *e* and *o*, they overgeneralize by interpreting that [ə] and [u], even when not alternating with those vowels, are correlated to *e* and *o*.

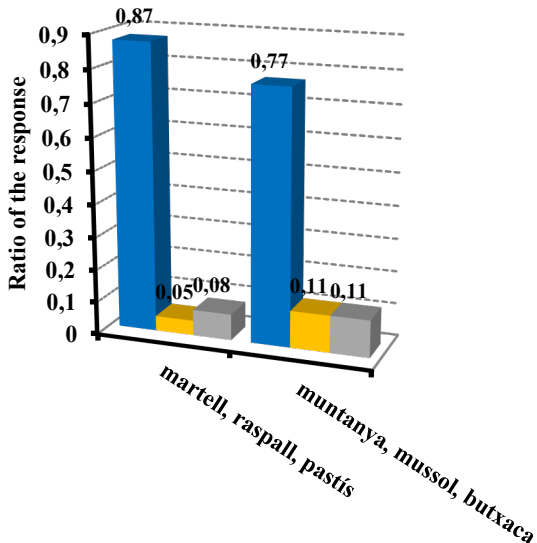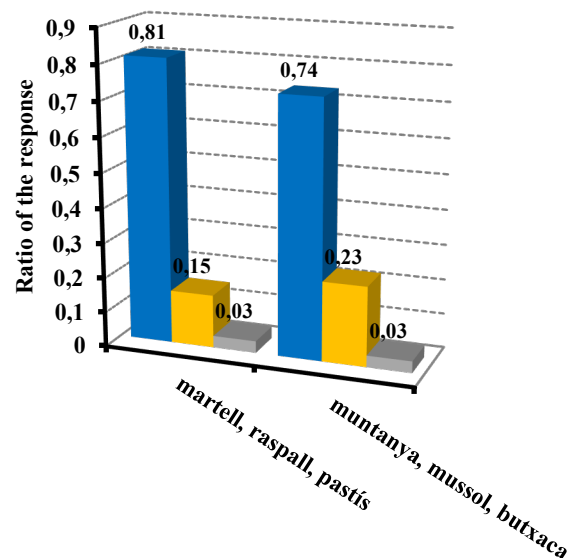FIG. 7. *Results for first grade children*   FIG. 8. *Results for second grade children*

7   Concluding remarks

This paper examines the written productions of first and second grade Catalan children in an attempt to tap into their understanding of the underlying forms of Catalan unstressed vowels for which alternations are not present (and for which multiple inputs are thus possible). The paper also explores the differences in spelling between alternating and non-alternating [ə] and [u], in order to find out whether morphophonemic alternations are taken into consideration in children's spellings choices, as well at which stage these start to be considered.

The results for first graders show generalized misspellings in which [ə] and [u] are systematically associated with the graphemes *a* and *u*, regardless these vowels alternate with vowels in stressed position or do not. This behavior clearly supports the Lexicon Optimization Hypothesis, according to which there is a first stage in language acquisition in which, from all possible input candidates (Richness of the Base), the learner selects the one that matches the adult output representation as the optimal input (Smolensky 1996).

The results for second graders show a noticeable decline in misspellings for alternating [ə] and [u]. This leads to think that morphophonemic alternations are progressively incorporated and taken into account in the process of UR construction. Of course, this better performance in the spelling of alternating vowels might also be a consequence of a better familiarity with conventional spelling. Importantly, though, the decline in misspellings for alternating [ə] and [u] coincides with an increase in spelling mistakes for non-alternating vowels. In our view, this undescores the influence of conventional orthography in children's spelling choices, points to the influence of phonology, and more specifically to a stage of vacillation with respect to the UR of non-alternating forms, with apparent overgeneralizations. This behavior, thus, may support the free-ride approach to morphophonemic learning (McCarthy 2005).

The results for both first graders and second graders show differences in the spelling between back vowels (with fewer mistakes) and non-back vowels (with more mistakes), and we argued this might be related to the fact that the children have to deal with a more reduced typology of output-input alternations for back vowels (two-to-one) than for non-back vowels (three-to-one). This is evidence, again, for the role of morphophonemic alternations (i.e. morphology) in children's spelling choices. The results for both graders also reveal that more productive alternations, such as the ones found between a base and the diminutive, are more transparent to the kids than others: whereas diminutives were generally spelled correctly, non-diminutives were not, giving support again to the influence of morphophonemic alternations, and their transparency, in children's spelling choices.

Traditionally, orthography has been regarded as being unrelated to phonological theory or to phonological knowledge. In this paper, though, we worked with the hypothesis that the way primary school children spell unstressed vowels may help to identify the underlying representation to which they associate non-alternating and alternating vowels. This hypothesis, which relies on the limited exposure of the children to conventional spelling at these stages, is borne out by the results of this study. Overall, we think children's spelling is a fair —and under-tapped— resource to shed light on our understanding of the process of acquisition and construction of underlying representations.

8   Samples of the responses

First grade

Second grade

9   References

Alarcos, Emilio. 1953. Sistema fonemàtic del català. *Estudis de lingüística catalana*. Barcelona: Ariel. p. 13-25.

Alarcos, Emilio. 1973. De fonología catalana: la vocal neutra. Archivum, 23. p. 293-297.

Badia, Antoni M. 1951. Gramática histórica catalana. Noguer: Barcelona.

Badia, Antoni M. 1965. Función significativa y diferencial de la vocal neutra en el catalán de Barcelona. Revista de Filología Española, 48. p. 79-93.

Bermúdez-Otero, Ricardo. 2006. Phonological change in Optimality Theory. In: Brown, Keith (ed.). *Encyclopedia of language and linguistics*, 2nd ed., vol. 9. Oxford: Elsevier. p. 497-50.

Bonet, Eulàlia & Lloret, Maria-Rosa. 1998. *Fonologia catalana*. Barcelona: Ariel.

Bourassa, Derrick C.; Treiman, Rebecca; Kessler, Brett (2006). Use of morphology in spelling by children with dyslexia and typically developing children. *Memory & Cognition*, 34. p. 703–714.

Fuertes, Gisela; Pons-Moll, Clàudia. 2021. On the UR construction of (non-) alternating unstressed schwas and [u]s in Catalan. Some evidence from primary school children handwriting. Talk presented at the 28th Manchester Phonology Meeting, Manchester.

Hamann, Silke; Colombo, Ilaria. 2017. A formal account of the interaction of orthography and perception: English intervocalic consonants borrowed into Italian. *Natural Language & Linguistic Theory* 35, 3: 683–714.

Krämer, Martin. 2012. *Underlying representations.* Cambridge: Cambridge University Press.

Maria-Rosa Lloret; Clàudia Pons-Moll (2016). Catalan Vowel Epenthesis as Evidence for the Free Ride Approach to Morphophonemic Learning. *Linguistic Inquiry*. Winter 2016, Vol. 47, No. 1. p. 147–157.

Mascaró, Joan. 1978. *Catalan Phonology and the Phonological Cycle*. Indiana: Indiana University Linguistics Club.

Mascaró, Joan. 1991. La importància lingüística de la polèmica sobre la vocal neutra del català central i el seu caràcter fonemàtic. *Els Marges. Revista de llengua i literatura*, 44. p. 33-43.

McCarthy, John J. 2005. Taking a free ride in morphophonemic learning. *Catalan Journal of Linguistics*, 4. p. 19-55.

Pons-Moll, Clàudia; Lloret, Maria-Rosa. 2013. About the underlying representation of words starting in VsC- in Catalan and the opaque character of morphophonological alternations by prefixation. Talk prsented at the 10th Old Word Conference in Phonology (OCP-10), Istanbul.

Pons-Moll, Clàudia; Lloret, Maria-Rosa. 2014. The free ride procedure to morphophonemic learning is correct. Some evidence from Catalan vowel epenthesis. Talk presented at the 22 Manchester Phonology Meeting, Manchester.

Pons-Moll, Clàudia. 2016. The limits of the free ride in UR discovery. Talk presented at the Phonological Theory Agora Meeting 3 ("Phonology and the Lexicon"), Tours.

Prince, Alan, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar.* Oxford: Blackwell.

Prince, Alan, Bruce Tesar. 2004. Learning phonotactic distributions. In: Kager, René; Pater, Joe; Zonneveld, Wim (ed.), *Fixing priorities: Constraints in phonological acquisition*. Cambridge: Cambridge University Press. p. 245–291.

Smolensky, Paul. 1996. On the Comprehension/Production Dilemma in Child Language. *Linguistic Inquiry*. Autumn 1996. vol. 27, No. 4. p. 720-731.

Treiman, Rebecca (2017). Learning to spell words: Findings, theories, and issues. *Scientific Studies of Reading, 21*(4), 265–276.

Treiman, Rebecca; Cassar, Marie; Zukowski, Andrea. 1994. What types of linguistic information do children use in spelling? The case of flaps. *Child Development*, 65. p. 1318–1337.

Varnhagen, Connie; McCallum, Michele; Burstow, Meridith. 1997. Is children's spelling naturally stage-like? *Reading and Writing: An Interdisciplinary Journal,* 9. p. 451–481.

Viaplana, Joaquim & Janet Ann, DeCesaris. 1984. La "vocal neutra" del català central: fonema o al·lòfon? *Estudis gramaticals*, 1. Bellaterra: UAB & Departament de Filologia Hispànica. p. 341-380.

Wheeler, Max. 1979. *Phonology of Catalan*. Blackwell: Oxford.

Wolter, Julie; Wood, Alexis; D'zatko, Kim W. 2009. The influence of morphological awareness on the literacy development of first-grade children. *Language, Speech, and Hearing Services in Schools*, 40, 286–298.