PRODUCTIVITY AND PORTUGUESE MORPHOLOGY: HOW EXPERIMENTS ENABLE HYPOTHESIS-TESTING

ANDREW NEVINS*
University College London

This paper examines to what extent, as evidenced by productivity studies, just as the robustness and learnability of phonological patterns may be grounded and biased by naturalness considerations, so may morphology. Four case studies (L-morphomes, competing nominalizing affixes, athematic participles, and defective verbs), that are carried out with experimental tasks using wug-words jointly demonstrate that choices in the realization of morphological categories may depend on submorphemic (features) and supramorphemic (phrasal syntax) principles, and show that both morphology-internal and interface considerations may modulate the extent of productivity of generalizations.

1. Introduction

The question to be examined here is the extent to which lexical trends and morphophonological generalizations interact with productivity. Once a learner has observed a trend within the formation of a certain morphological category, to what extent do they generalize it beyond the static lexicon? This question has been extensively asked for phonological patterns of varying types (e.g. phonotactics and phonological processes / constraints; see Zimmer 1969, Berent et. al 2009, Finley 2012, Becker, Nevins & Ketrez 2011 among many, many others); for example: if a word has a round vowel in the first syllable, will it automatically disallow a non-round vowel within the second syllable, and do native speakers pick up on this to the extent that they have well-formedness judgements about non-existing novel words? If a language shows evidence that high vowels and voiceless consonants do not occur word-finally, is that too obscure to generalize? Do learners have a good set of 'hunches' about what kind of constraints are ones that, no matter how very well may hold within the lexicon, still look too accidental to enforce in wug-words?

* This paper is an overview and synthesis of four different studies within the same overall line of research, all of which are joint research projects with the invaluable collaborations of my coauthors, as listed in the respective sections. I thank Paula Armelin, Clàudia Pons-Moll, Ezekiel Panitz, and Joanna Zaleska for discussion of the present paper, as well as the audiences, organizers, and reviewers from Going Romance (Amsterdam 2013) and NYU Abu Dhabi.

Within the realm of a predictive morphological generalization, questions of these sort have received less attention, although they are very much ripe for the asking. For example, suppose that across the lexicon of verbs, the form of the second singular is predictable from the form of the first singular (as opposed to any other form within the inflectional paradigm) – is this reasonable enough to be applied with confidence to novel words for which the entire inflected paradigm is not known? What about, say, forming denominal verbs, given a choice of two suffixes, M and N. Will the choice between M and N, even if it depends on a potentially obscure fact about whether the word ends in a mid-vowel or not, be salient enough and 'make enough morphological sense' to leverage in an experimental task when deeming the well-formedness of two potential competing candidates?

In a sense, the unparallel trajectories of these two areas of study through experimental means reflects the state of development within their theoretical formalizations as well. (It is arguably a truism that the extent to which experiments of this sort will be revealing is in direct accordance to the degree to which they have correspondence with a set of predictions stemming from particular theories). Asking a phonologist if a pattern of seemingly pathological palatalization exists or not according to a particular set of theories of phonological *naturalness* is pretty straightforward these days, as a voluminous set of substantively-biased and formally biased predictions have been formulated (Hayes, Steriade & Kirchner 2003, Moreton 2008, Heinz 2010). However, the parallel trajectory of 'Natural Phonology' in the 1970s (reflourishing in the 1990s with a host of increased dialogue between phonological typology and phonetic research) has not proceeded alongside much consensus as to what constitutes Natural Morphology, despite occasional valiant efforts in the 1980s largely limited to Central Europe (e.g. Dressler, Mayerthaler, Panagl & Wurzel 1987). The research in this area has been overall too sparse (and often too disconnected from some of the most direct interface areas with morphology) for a real consensus anything like what we have in phonology today to crystallize yet.

The dominant trend in many corners of morphological research is to say 'Nothing is natural'. Statements of this sort express a degree of nihilism for predictive morphological typology to an extent that doesn't even parallel Anderson's (1981), Blevins' (2004) or Hale & Reiss' (2008) eschewal of any systematic crosslinguistic patterns – as these authors directly embrace formal simplicity as an evaluation metric and thereby a predictive measure of the overall naturalness, acquirability, diachronic persistence, and productivity of certain logically possible patterns. Take one core tendency of morphological inflection: defectiveness in verbal or nominal paradigms, a phenomenon found time and again in languages as distinct as Russian, Mohawk, and the Iberian languages. Focusing on the latter of these, Maiden & O'Neill (2010: 112) state that "The major domain of defectiveness (present subjunctive and 1sg present indicative) seems, and is, irreducibly

arbitrary if one seeks a motivation outside the morphological system itself." (emphasis added).

Their conception of limits on possible morphological patterns (found more generally within 'autonomous' approaches to morphology) is that morphology needn't answer to phonology or syntax; that it follows its own logic, however bizarre it may seem. On this view, learners are directly endowed with the ability to sift through and compute morphological productivity (or lack thereof, in the case of defectiveness) purely by problem-solving abilities exclusive to morphological geometries; a kind of sudoku puzzle to be eventually mastered given enough internal and external reward in the learning process; a poker hand dealt by the exigencies of diachronic drift, in which an L-shaped pattern (borrowing a metaphor from the Roman alphabet) can be learned in a two-by-three geometric paradigm (e.g. all [-plural] cells and the [+plural, -participant] cell, but not the [+participant] plurals). Is such a pattern easier or harder to learn – i.e. more natural -- than a **п**-shaped, ד-shaped, or и-shaped pattern? The right answer is not yet in sight, and can come from many directions – among them extensive crosslinguistic typology in which database comparisons inform patterns of natural vs. unnatural syncretism and formal modeling of the computational primitives that favor one system over another (e.g Bobaljik 2012, Pertsova 2011, Graf 2012), and in the case of the present paper, adopting the tack of creating closely-controlled experimental conditions in which native speakers must choose between one of two potential, non-existing candidates, or rate the well-formedness of unseen inflected forms of wug words. In the current paper, we examine four case studies in Portuguese, within hopes of contributing to the above triangulation between theoretical modeling, typological sampling, and experimental comparison of the relative naturalness of certain generalizations learners may adopt or neglect.

Work within many (but not all schools) of morphological theory within the past half-century have proposed certain formal primitives, substantive constraints, and interface-derived properties that systems of inflectional and derivational morphology 'like to have'. (One might cartoonishly say that such efforts embody the 'Prague School' approach to morphology, as opposed to the 'Hyper-Saussurean' extension of arbitrariness to even closed-class morphological formatives embraced by Maiden, Aronoff, and others). Among the components in the potential naturalness toolkit investigated herein are:

> (1) *Featural decomposition* – that an inflectional paradigm is constituted by 'natural classes', akin to the ones familiar from segmental phonology, that are directly formed by a small set of, say, binary-valued features that characterize particular 'rows' or 'columns' of the paradigm, as opposed to others. In fact, many distinct systems of featural decompositions are possible, and research often

employs relative attestedness vs. unattestedness as a means of comparison among them, along with a generalized dispreference for disjunctive statements (e.g. the equivalent to 'all voiced obstruents and non-lateral, non-glide sonorants) instead of concisely-defined natural classes (e.g. [-continuant]) in morphological terms).

(2) *Markedness-based neutralization* – that within a paradigm of oppositions (e.g. comparative versus more marked superlative, dual versus more marked plural, or first conjugation versus more marked second and third), not all of them will show an equally parallel number of subdistinctions, but rather the marked ones will favor (and perhaps actively strive towards) syncretism more than others.

(3) *Structure-Based Spellout* – that morphological formatives (and phenomena conditioning their distribution, such as syncretism, allomorphy/suppletion, and defectiveness) at times demonstrate a direct and clear grammatical sensitivity to the details of the internal hierarchical structure responsible for the morphosyntactic properties of the item (thereby in direct contrast to models such as Anderson 1992, in which the hierarchical morphosyntactic structure of words *cannot* influence any aspects of its morphophonological form).

These elements of morphological naturalness, and the ways in which they generate direct predictions about preferred and dispreferred learning-and-generalization patterns, are investigated here with a focus on Romance languages, primarily Portuguese (its European and Brazilian variants), with relevant comparison to Spanish and Italian, and explore well-understood phenomena within these languages that heretofore have largely gone unexplored in terms of experimental morphology. The present article in part is a synthesis and review of research reported in greater detail elsewhere, reflecting a number of collaborative efforts cited within the subsections that follow.

Study 1, L-Shaped Morphomes (§2): A Productive case of Lawless Syncretism? (Aronoff 1994, Maiden 2005). Answer: no, learners actually refuse to generalize this pattern to novel verbs, and opt for one that 'makes inflectional sense' instead, when given a choice.

Study 2, Nominalizing Affix choice (§3): Given a choice between suffixes M and N, is affix rivalry memorized, verb-by-verb, with competition due to language-independent "principles of self-organizing systems"? (Lindsay & Aronoff 2010)? Answer: no, learners opportunistically choose to extend affix choice among two nominalizers directly insofar as it increases the 'reach' of a pattern of morphophonological neutralization already nascent elsewhere in the morphology of this specific language.

Study 3, Athematic Participles and the 1sg again (§4): Has a completely accidental pattern of syncretism, this time involving forms not even within

the same tense/aspect paradigm, been reinterpreted as arbitrary-but-systematic, since it seems that short-forms are preferred to long-forms and something must govern this choice, however arbitrary it seems (Chagas 2007)? Answer: no, the choice of short and long-forms in participles seems to directly reflect the hierarchically-built complexity of their argument structure.
Study 4, Defective Verbs in Portuguese (§5): Are no generalizations possible outside the morphological system itself as to why and when forms will be defective? Answer: in fact, it is largely predictable based on the phonology of the verb stem whether it is a candidate for defectiveness or not.

The studies largely underscore the importance of morphology not-by-itself, but rather, paralleling 'grounded phonology', firmly informed by properties of the interfaces with syntactic structure, phonological patterning, and internal patterns of asymmetrically-organized featural decomposition.

2. L-Shaped Morphomes

The L-shaped morphome in Romance (Maiden 2005) refers to the fact that the stem in the first person singular of the present indicative (henceforth 1sg.ind, e.g. *dig*-o in (4) and *ouç*-o in (5) is systematically identical with the the stem in entire set of present subjunctive forms, e.g. the 1sg.sbj *dig*-a and the 2sg.sbj *dig*-as) in terms of the stem *dig*- (as opposed to *diz*-), or *ouç*- (as opposed to *ouv*-).

(4) dizer 'to say'

	IND	SBJ
1sg	dig-o	dig-a
2sg	diz-es	dig-as
3sg	diz-Ø	dig-a

(5) ouvir 'to hear'

	IND	SBJ
1sg	ouç-o	ouç-a
2sg	ouv-es	ouç-as
3sg	ouv-e	ouç-a

This is not simply like having irregular verbs *bring* and *brought* in English: what is irregular here is not only the phonological relation between the forms themselves (opaque/mysterious in terms of synchronic segmental processes, just like *bring/brought*), but more importantly their distribution:

(6) a. Why should the 1sg.ind show identity with the 2sg.sbj, while the 2sg.ind does not?

b. Why should the 1sg.sbj and 2sg.sbj show identity with each other, while in the indicative, no such identity is found?

According to Maiden (2005) this L-shaped relation among unrelated forms is what constitutes a 'morphome', and shows the productive autonomy of the morphological component by itself: it needn't answer to phonology or syntax. The rule was in fact originally phonologically motivated within verbs from the *–er* and *–ir* class (the second and third conjugations), having to do with the fact that both the 1sg.ind and all of the sbj forms had a [+back] vowel following them, while the others had a [-back] vowel, at a time during which palatalization by front vowels was active in the language. However, long after palatalization ceased to be productive, the pattern was extended to forms with no such conditioning possible to state, e.g. *ouç/ouv*, phonetically [ows/owv]. Moreover, from a featural perspective, there is no reason these should form a natural class; they are disjunctive. The question, therefore is whether such patterns are synchronically productive within the Romance languages, a question all the more important to ask as the L-shaped pattern is *trivially true* for even all <u>regular</u> verbs in the language:

(7) falar 'to speak'

	IND	SBJ
1sg	fal-o	fal-e
2sg	fal-as	fal-es
3sg	fal-a	fal-e

It is the case that the 1sg.ind and the sbj forms, e.g. 2sg.sbj are in fact identical to each other in *fal-/fal-* in (7); if the learner includes regular verbs within the scope of their generalization, the L-shaped pattern potentially holds of every verb in the language.

2.1 Experimental Investigation

The basic experimental design here is to present speakers with an 'incomplete' paradigm and ask for the 2sg subjunctive. More specifically, imagine a paradigm of six cells, in which a participant is presented with only two of them, say C1 with form *nep-* and C2 with form *nef-*, each presented within appropriate morphosyntactic contexts. The participant is then asked to provide the form for what the morphosyntax of the frame sentence (e.g. the tense, mood, and features of its pronominal subject) would require as C3. Whether they consistently choose C1 or C2 as the stem on which to base the unseen form – in principle a coin flip 50%-50% choice all else being equal – potentially bears on whether the L-shape is alive and kicking or not. If *morphomes* – that is, morphological outlaws that do not correspond to the notion of *morpheme*, as they represent disjunctive distributions with no grounding in morphosyntax or phonology – have any

explanatory or ontological force, then non-interface-grounded distributions should be chosen. After all, speakers of European Portuguese (with a healthy ongoing use of the subjunctive mood) have abundant evidence staring them in the face that irregular verbs in their language show such 'diagonal syncretisms'.

We set things up so that speakers were provided with the 1sg.ind and the 2sg.ind, and probed for the 2sg.sbj, as shown in the sample frame in (8).

(8) Eu *nepo* muito rápido, e tu também *nefes,* embora tu ainda não _____ rápido.
'I *nepo*$_{1SG.IND}$ very quickly, and you *nefes*$_{2SG.IND}$ too, even though you don't _____$_{2SG.SBJ}$ quickly yet.'

Given this incomplete paradigm, their choice of a base for the 2sg.sbj, can be either based on the L-shaped morphome (in which case it would be the 1sg.sbj stem *nep*-), or instead on choosing the stem for 2nd person across moods, whether indicative or subjunctive (in which case it would be *nef*-). These two choices were made to be mutually exclusive, and participants' responses for the 2sg.sbj were classified into 1sg.ind base or 2sg.sbj base. Details that cannot be included within scope of the present chapter are to be found in Nevins & Rodrigues (2012), to which the reader is referred for a more extended report and discussion. As shown in Figure 1, the results of this experiment, conducted with n = 250 participants, revealed that 67% of the time, rather than going for the unnatural, diagonal syncretism of the L-shaped pattern, responses instead showed a preference for the same stem across moods within the second person (i.e. a 'horizontal' syncretism, given the arrangement of the tables in (4) & (5)).

In other words, the L-shaped pattern appears to be no longer productive (and hence no longer part of the morphological grammar of native speakers of European Portuguese), as they systematically rejected the opportunity to extend the distribution of stem choice in accordance with the pattern nonetheless present in the language. (It would be interesting to address in future studies whether the 33% responses that *did* show a preference for the L-shaped pattern come from speakers with a higher degree of metalinguistic or prescriptive awareness). By and large, these participants thus manifested a naturalness bias in which, given the chance to base syncretism on morphosyntactic natural classes expressed by featural decomposition, instead of based on cumbersome disjunctive statements (see Bourne 1970, Feldman 2000 for evidence that these are cognitively more challenging to learn), they indeed chose the former.

When we initially presented these results at a workshop in Coimbra in 2011, Martin Maiden (pers. comm) responded that the failure to find evidence for the L-shaped pattern being productively extended in Portuguese might be due to the fact that, unlike the other Romance languages Spanish and Italian, in Portuguese, there are fewer consonantal

alternants showing an L-shaped distribution (many such verbs, e.g. *conosk- / conos-* 'to know (someone)' in Spanish having been restructured in Portuguese to be single-stem verbs). While true that Portuguese shows fewer consonantally-based L-shaped morphemes, it does show even more vocalically based L-shaped morphemes, e.g. *fugir:*

(9) fugir 'to flee'

	IND	SBJ
1sg	fuj-o	fuj-a
2sg	fog-es	fuj-as
3sg	fog-e	fuj-a

Nonetheless, one could maintain that unnatural inflectional identity is more salient and systematic for consonants than vowels (see Steriade 2008 for evidence in this direction based on Romanian). We thus set out to directly test both aspects of this objection, by replicating the experiment in Italian, with both consonantal and vocalic alternations showing the L-shape. However, yet again, an experiment with a large number of participants (n=175), presented in Figure 2 showed an overwhelming preference for basing the 2sg.sbj on the 2sg.ind, rather than its L-shaped congener, regardless of the consonantal vs. vocalic locus of the alternation.

A further potential objection, helpfully raised by various audience members in Newcastle and York, was that while the results of these first two experiments show that there is indeed a preference for horizontal (e.g. person-based) syncretism as opposed to arbitrary L-shaped syncretism, nonetheless it may be the case that the L-shaped pattern *is* learned and part of the grammar, but is simply overshadowed by a stronger preference for consistency-of-person. We addressed this in an additional replication, this time conducted in Spanish, the frames for which are shown in Figure 3. This time, participants were provided with the 1sg.ind and the 2sg.ind, and probed for the 3sg.sbj. In this case, given this incomplete paradigm, their choice of a base for the 2sg.sbj, can be either based on the L-shaped morphome, or on the 2sg.ind, a choice which is not directly reducible to 'consistency of person across moods'. As such, the experiment design gave the L-morphome its best chance to flourish. Nonetheless, as the results show (n = 148), participants *still* opted for the 2sg.ind as a base, thereby consistently refusing the L-morphome pattern. Their choice, instead, seems to be based on *avoiding* the 1sg.ind as a base for derivation, perhaps as it is known to be an 'unreliable' source – in more formal terms, this avoidance can be cast as a preference for choosing the same inflectional base across persons, but at a more fine-grained level, namely among [-author] persons.

Taken together, the results of these experiments conducted across three different Romance languages with a variety of experimental manipulations and large numbers of participants demonstrate that the L-shaped morphome is no longer productively extended to the subjunctive. This would in turn

support a view of learning biases in morphology, whereby unnatural and disjunctive patterns are biased against, and only can be learned with overwhelming amounts of evidence to the learner that such biases should be overturned. While proposals about the existence of 'morphomes' merit extremely strong evidence (just as proposed phenomena in phonology which counter 'natural' trends do), and while we have demonstrated above that the L-morphomes of Romance seem to have little to no synchronic reality, these patterns did arise in the language at some point in their history. As Maiden (2005) shows, the L-morphome pattern *was* productive around 600 years ago, as it involved cases in which historically non-morphomic verbs acquired this disjunctive distribution of stems. Accepting for the sake of argument that the L-morphomes were semi-productive 600 years ago (perhaps given overwhelming relative evidence to the learner) our experimental results demonstrate that something has changed since then. In the next subsection, we conducted a historical corpus study in order to examine what might have led to the decline of the L-morphome.

2.2 Towards a Historical Explanation of the Loss of the L-Morphome

The L-shaped morphome *was* productive 600 years ago (Maiden 2005), when new verbs were ushered into this pattern. What changed since then? The answer we wish to explore is, at first blush, deceptively simple: there are simply more verbs in the Romance languages now than there were in the fifteenth century. Of course, having more verbs alone is not enough – what seems to have crucially changed is the *ratio* of 1st conjugation verbs to 2nd/3rd conjugation verbs. While 1st conjugation verbs have always been more productive, the difference between them and the 2nd/3rd conjugation verbs has become more extreme over time – rather than having stayed at a stable ratio of, say, 55% 1st conjugation over time, successive generations of learners have allowed the rich to get richer, and the 1st conjugation to grow in proportion over time at a rate beyond that of the 2nd/3rd conjugation. We conducted corpus studies in Portuguese, Spanish, and Italian, and found that for all three of these languages, the proportion of 1st conjugation verbs to 2nd/3rd conjugation verbs has increased overtime. As a result, the salience of the L-morphomic patterns found within the 2nd/3rd conjugation have ceased to receive the amount of lexical support they need to survive, particularly in the face of the fact that it is an unnatural distribution to begin with. Details of these corpus studies that cannot be included within scope of the present chapter are to be found in Tang & Nevins (2013), to which the reader is referred for a more extended report and discussion.

As shown in Figure 4, in the period from 1525-1899 for European Portuguese, not only has the number of verb types grown over time with fixed vocabulary samples, but the ratio of *-ar* to *-er/-ir* has increased as well, as indicated by the 'mean productivity ratio'. Importantly, this study has been conducted with a *fixed* number of samples per epoch – it is not the

case that our findings are simply the results of the fact that there are more texts available in more recent periods. As the sample size is fixed across epochs, but nonetheless the ratio of types *within* these equal sample sizes, the results thus demonstrate a growing disparity between the ratio of verbs in the first conjugation relative to those of the second and third conjugation since the 16th century. Similar results are presented for Italian from 1550-1974 in Figure 5 and Spanish from 1522-1996 in Figure 6. Orthogonally to the question at hand, we observed the largest and most consistent sudden increase in verb vocabulary size (therefore productivity) at around 1750 across the three Romance languages, as confirmed by an objective changepoint statistical analysis. It is tempting to speculate that that the reason for this sudden jump in the diversity of types in the lexicons of these languages is a result of the technological, medical, and educational innovations that followed in the wake of the Industrial Revolution.

Whatever the socioeconomic, cultural, and historical reasons for the changes across the lexicon of verb types over these nearly six hundred years, the result is a decreased representation of the 2nd and 3rd conjugations within the lexicon – precisely the conjugations in which the irregular and unnatural morphomic patterns reside. As a consequence, if indeed learners need increased evidence in order to incorporate and actively uptake unnatural patterns, this lexical support has dwindled over time. Even though many of the morphomic verbs have maintained a very high token frequency (allowing them to survive as memorized), their type frequency has diminished over time, and hence they go unlearned as a generalizable pattern. When the distribution of irregular alternations is overshadowed in the lexicon (cf. Yang 2005), a morphologically unnatural pattern may cease to be productive.

3. Affix Choice Reflects Bias Towards Conjugation-Identity

The two verbal nominalizers in Brazilian Portuguese *–ção* and *–mento* show a competition somewhat similar to *–ity* and *–ness* in English (cf. Anshen & Aronoff 1997). While some doublets exist, e.g. *afinação/afinamento* 'tuning', from the verb *afinar* (1st conjugation), even for these forms, the *–ção* variant is often preferred. However, while *–ção* is vastly more productive than *–mento* across 2000 verb forms in the Houaiss dictionary (see Figure 7), verbs of the second conjugation show an unusual reversal in their preference for *–mento*. This is all the more surprising as *-mento* causes a particular morphophonological change in the theme vowel: *receber / recebimento*. This affix-specific change, however, enforces a neutralization of the theme vowel between the 2nd/3rd conjugation in derivational morphology, found as a general result with 2nd conjugation verbs, e.g. *correr* 'to run', *corrido* (participle).

We will compare herein two hypotheses to account for the choice of *–ção* and *–mento* in deverbal nominalization. Under the first of these,

speakers must memorize which nominalizer goes with which verb; conjugational trends are simply effects of 'self-organizing systems'. An alternative, however, is that *–ção* is indeed more productive; the default nominalizing affix across the language. The reversal of the trend in the 2nd conjugation is specifically related to the morphological rule of vowel-raising that *-mento* triggers. According to Oltra-Massuet 1999: the 2nd and 3rd conjugation form a marked subgroup, which can be and are often neutralized (not only in the deverbal formations cited above, but also in their choice of theme vowel in the subjunctive, in which these two conjugations are neutralized).

As for why vowel raising only applies with *–mento*, we adopt the proposal of Freitas (2014), that it spells out in the same domain as its verbal base, while *–ção* does not . The basic intuition is that *–mento* is closer to the root + theme vowel (i.e. the stem) than *–ção* (a difference which may reflect either Lexical Phonology style strata of affixation, or differences in the height of syntactic attachment). In Freitas' analysis, *–mento* and the stem are spelled out in the same cycle, whereas *–ção* is spelled out in a cycle after the stem. As a result, *–mento* but not *–ção* has the possibility of inducing raising of the theme vowel in second conjugation verbs to [i], thereby effecting a neutralization between the 2nd and 3rd conjugations.

The experimental research we conducted, therefore, set out to test the following question: is the preference for *-mento* in wug-words from the 2nd conjugation simply the result of committing to memory an arbitrary pattern? Our hypothesis is that the 'reversal' of the preference in the 2nd conjugation is a) directly encoded by speakers as a generalization to be extended to novel verbs, and b) a direct consequence of the morphophonological rule of theme-vowel neutralization it induces. In order to examine these hypotheses, we conducted two experiments. In the first, we simply asked speakers which of the formations, *-mento* or *–ção*, they preferred for novel verbs across all three conjugations; a sample test item is shown in (10)

(10) *Pedro tem mania de treter. Sempre que viaja, ele compra algo novo para aumentar { sua treteção / seu tretimento}*
'Pedro has a compulsion to *'treter'*. Every time he travels, he buys something new to increase his 'tretification' / 'tretiment'.'

In the second experiment, we included a 'pathological' version of *–mento*, one which did not incur vowel raising.

(11) *As irmãs da Joana não gostam de penzer. Elas sempre se queixam para sua mãe como é ruim { o penzemento / a penzeção}*.
'Joana's sisters dislike having to 'penzer'. They always complain to their mother how awful they find 'penziment' / 'penzification'.'

Details that cannot be included within scope of the present chapter are to be

found in Freitas & Nevins (in preparation), to which the reader is referred for a more extended report and discussion.

As shown in Figure 8, the results of the first experiment (n=140) demonstrate that speakers are aware of the conjugation-specific preference for *–mento* over *–ção* only in the second conjugation, with a preference for *–ção* otherwise. More strikingly, in Figure 9, the results of the second experiment (n=103) show that the difference between the 2nd and 3rd conjugation disappear as soon as the vowel-raising rule is removed from the equation. Suppose that the affix choice, therefore, is modeled in terms of two ranked constrants, informally presented as follows: WHEN DEVERBAL, DON'T DISTINGUISH THEME VOWELS I/E >> PREFER –ÇÃO OVER –MENTO. The presence of a vowel-raising *–mento* (e.g. with a floating [+high] feature) means the former constraint will trump the latter in the second conjugation, but when the pathological [+high]-less *–mento* is introduced, the former constraint is silent, and the preference reverts to the default *-ção* with the latter constraint.

This experiment is crucial in showing that the pairing of 2nd conjugation verbs with the affix *–mento* is <u>not</u> only about the affix, but rather about the vowel-raising rule it triggers, since as soon as the latter is removed, but the affix remains, the preference disappears. The results are akin to a 'placebo' comparison in a certain sense: they show the advantage of using experimental means in order to 'artificially' decouple two things that co-occur in 'the natural world', and after removing only one of them in the laboratory, examine whether the other still exerts its effects in order to deduce the causal structure of the whole.

4. Athematic Participle Choice Predicted by Argument Structure

An interesting current development in Brazilian Portuguese (BP) involves the short form passive and perfect participles that don't look like the corresponding infinitive, as in the passive and perfect in (12b), the perfect in (12c), and optionally the passive and the perfect in (12d).

(12)

infinitive	1sg past	1sg pres	passive	perfect
a. apagar 'extinguish'	apaguei	apago	ser apagado	ter apagado
b. acender 'ignite'	acendi	acendo	ser aceso	ter aceso
c.frigir 'fry'	frigi	frijo	ser frigido	ter frito

d. comprar ‘buy’	comprei	compro	ser comprado / ser compro	ter comprado / ter compro

Forms like (12b) *acender/aceso* are largely frozen, though examples of this sort are widely found in Italian (e.g. *mettere/messo* ‘to put’); throughout the Romance languages most of these reflect diachronic developments from the so-called *t*-stems inherited from Latin (see Bachrach & Nevins 2008 for an overview of approaches to this phenomenon). Notably, these irregular participles can be called *athematic*, as they lack the theme vowel (preserved in *apag**a**do* but lost in *aceso*). For this reason, they are sometimes called ‘short-form’ participles; in Lobato (1999), the analysis is that these innovative participles involve an alternative spell-out, without the theme vowel and the *-d-* morpheme (the nature of which has seen various proposals across the Romance languages, e.g. deverbalizer, resultative, etc.)

What is unique to contemporary BP is the emergence of athematic participles that do not reflect this inheritance, and indeed, some of which are considered substandard or variant, but are nonetheless increasing in attestation. The forms in question in fact show doublets, thus *ter comprado* and *ter compro* are both acceptable, thereby standing in marked contrast to the more entrenched cases such as *ter frito*, alongside which **ter frigido* is unacceptable (note also that the innovative *ter trazido / trago* ‘to have brought’ and innovative *ter perdido / perco* ‘to have lost’ are both acceptable, in contrast to the entrenched ones). Interestingly, for some verbs, the passive particle favors the short-form while the perfect favors the long form (e.g. imprimir ‘to print’, with *ser impresso, ter imprimido*), while in others, this preference is reversed. While this potential difference between the passive and perfect participle merits further exploration, our focus was to test a range of verb types, including unaccusatives (which cannot be passivized), and hence we limit our attention to the perfect participles in what follows.

Under analyses such as Lobato’s, the short-forms are alternative spell-outs of a syntactic structure, and no matter which details of that structure one may conclude, the intuition is that shorter forms reflect lesser structure. As such, the specifics of the hierarchical internal structure of words comes to be relevant in understanding their morphophonological form. Note that such a mapping principle is not present in all approaches to morphological realization, even ones that share the assumption of ‘late insertion’ adopted in some form or another in work such as Lobato’s (cf. Embick (2004) for related analyses in English).

In fact, as (12) reveals, one tempting generalization that can be put forth is that the innovative participial forms are identical to our old friend, the 1sg.ind. Now while this is not the case for the entrenched forms in (12a-

c), Chagas (2007) puts forth the view that the relation between the 1sg.ind and the athematic participles is increasing in productivity. (Even though, as should be evident by now, the 1sg.ind in particular has no privileged morphosyntactic closeness to participial forms that would not be shared by other inflected person/mood combinations). In other words, are we dealing with another kind of morphome? Well, to be a morphome, it must be productive, and have no relation to the phonology *or* the syntax. Conversely speaking, a truly arbitrarily learned inflectional mapping between 1sg.ind and short form participles resulting from 'purely morphological correspondences' should show no sensitivity to syntactic properties of the verbs in question, as long as they are all put in the same participial environment.

By contrast, under a theory in which the relation between short-form participle spellout and the word-internal syntax of participial expressions is directly reflectable, properties such as argument structure and event structure can play potential roles in conditioning the acceptability of such forms. Thus, it is relevant to mention that of the extant innovative short-forms, very few are found with ditransitives and psych-verbs such as 'to please', 'to annoy', 'to fear'. Nonetheless, this may be related to token frequency in the language or other ungodly factors. For this reason, wug-testing is extremely useful, as all of the forms presented have a token frequency of zero. Participants can in principle rely on the phonological form, the syntactic environment, or neither of the above, in rating the goodness of short-form participles given a particular infinitive and 1sg.ind form. Thus, in the experiment we conducted, we presented frames as shown below in which the infinitive (e.g. *pandar*) was presented in the first clause, and the 1sg.ind (e.g. *pando)* in a later clause, before eliciting a question of the well-formedness of short vs. long forms in combination with the perfect auxiliary.

> (13) Eu acho que *pandar* a casa todos os dias é importante para manter tudo limpo. Eu *pando* a minha casa bem, mas minha mãe, que é mineira, tem um capricho incrível e *panda* muito bem. No ano novo, depois que ela tinha (*pando/pandado*) toda a casa, ficou um cheiro delicioso.
> 'I think that to 'pandar' the house every day is important in order to keep everything clean. I 'pando' my house very well, but my mom, who is from Minas, has an incredible knack and 'panda' very well. On New Year's Eve, after she had 'pando/pandado' the house, it smelled wonderful.'
>
> (14) *Malir* dinheiro para familiares é muito perigoso. Eu nunca *malo* grana para ninguém, mas o Fernando vai ficar na miséria. Ele *male* toda a grana que ele tem para os primos. Se eu fosse ele, eu não tinha (*malido/malo*) 40,000 Reais para aquele primo esquisito dele que

mora na Alemanha.
'To 'malir' money to relatives can be dangerous. I never 'malo' funds to anyone, but poor Fernando might end up penniless. He 'males' all his money to his cousins. If I were him, I wouldn't have 'malo/malido' 40,000 bucks to that weird cousin of his living in Germany.

Crucial to the experimental design was an equal number of trials in four different verb classes: unaccusative, psych-verb, ditransitive, transitive. Filler trials involved unrelated morphophonological alternations such as the diminutive, *-ção* vs. *-mento*, and other nominally-based productivity-related alternations. Details that cannot be included within scope of the present chapter are to be found in Nevins & Rodrigues (2014), to which the reader is referred for a more extended report and discussion.

The results (n=98) are shown in Figure 10. Unaccusatives displayed the highest acceptability for short forms, followed by transitives and ditransitives (the difference among which was non-significant, suggesting they are treated the same), in turn followed by psych-verbs, these last of which showed the least acceptance of short-forms. These results demonstrate a clear sensitivity to syntactic aspects of the sentences in which the words in question were presented; in other words, the morphological form of these participles cannot remain isolatedly staring inwards; speakers pay attention to phrasal aspects of the verb in question in determining whether a theme vowel and the *–d-* are omissible or not.

A straightforward interpretation of these facts emerges rather easily, once the literature on the verbal classes is considered. Specifically, under the hypothesis that unaccusatives have no little *v*, (di)transitives have a little *v,* and psych verbs little *v* plus Causer (e.g. Pesetsky 1995), the ranking of the three levels of omissibility in short-forms directly corresponds to the amount of structure: those verbs with the least amount of heads tolerate short-forms the most, while those with a richer structure prefer to expone it using the long form. Suppressability of *-d-* thus corresponds to amount of structure; this parallels a great deal of existing work on the way that morphological marking reflects additional syntactic structure building, in a way potentially related to the Monotonicity Hypothesis of Koontz-Garboden (2008).

The relevance of structure may not be limited to simply counting the number of arguments (although this works surprisingly well); in addition, event-structural properties of the verbs (as diagnosed by the relative admissibility of telic-oriented adverbs such as *in two hours* and *get*-passives) pair along with the well-formedness of short-form passives in the set below.

(15)
a. Unaccusative: The vase broke *for/in two hours

b. Unaccusative: The vase got broken
c. Ditransitive: John sent the document to Mary in/?for two hours
d. Ditransitive: The document got sent to Mary
e. Transitive: John cleaned the house in/for two hours
f. Transitive: The house got cleaned
g. Psych-verb: John feared the situation *in/for two hours
h. Psych-verb: *John got feared

In future work, we plan to look at the distinction between specific types of psych-verbs, to tease apart theories in which subject-experiencer and object-experiencer verbs may have different amount of structure. The present results point to the fact that morphological realization *can* make reference to the amount of hierarchical syntactic structure underlying it (*pace* models such as Anderson 1992). In sum, suppression of the theme vowel and *-d-* in the short-form participles is not simply a morphomic generalization based on analogical reasoning stemming from Latin *t*-stems (*acendido/aceso*). Instead, speakers reliably show tolerance for short-forms when they correspond to smaller amounts of argument structure.

5. Defective Verbs and Postma's Generalization

Spanish and Portuguese are well-known for their 'defective' verbs: those which possess an infinitive, participle, and preterit forms, but a limited set of present tense forms. One such Portuguese example is *abolir* 'to abolish', which possesses the infinitive, the participle and gerundive *abolido* and *abolindo*, the full set of preterit and future forms, but no present tense indicative except for the 1pl *abolimos,* and no present tense subjunctive forms whatsoever. A notable generalization about the 'missing' forms is that they are in a large part, 'rhizotonic', i.e. they would have stress on the root, rather than on an inflectional desinence. (Correspondingly, it may be for this reason that Italian and Catalan, which allow rhizotonic infinitives, seem to have no defective verbs). Nonetheless, this generalization is not the whole story, as the would-be 1pl subjunctive form *abolamos* would bear stress on the penultimate syllable (and not the root), and thereby the correct generalization seems to be that the only 'effable' forms are those which overtly retain the theme vowel /i/ after the root. The preceding statement thus implicitly carries one of the generalizations at hand: defective verbs in Portuguese are all in the 3rd conjugation.

Characterizing the set of verbs which show this defectivity (as clearly, not all 3rd conjugation verbs are defective) has proved an elusive task for many grammarians. While some of them might be classified in terms of semantic restrictions (e.g. the putatively defective verb *latir* 'to bark' would sound odd or uncivilized in the 1sg), the fact that such forms are also defective in say, the 3rd person subjunctive rules out any appeal to semantics alone. In characterizing the set of defective verbs, an important caveat is

also to make sure that speakers *know* the verbs at all; those which are so low-familiarity as to show uncertainty about any of the forms should be excluded from the first pass of arriving at a generalization. One of the often tempting explanations is based on homophony avoidance, as the three defective verbs below would have a 1sg indicative and subjunctive identical to that of another, non-3rd conjugation, and high-familiarity verb:

(16)
a. polir 'polish': 1sg.ind *pulo,* 1sg.sbj *pula* (forms confusable with *pular* 'jump')
b. parir 'give birth': 1sg.ind *paro,* 1sg.sbj *para* (forms confusable with *parar* 'stop')
c. falir 'fail': 1sg.ind *falo,* 1sg.sbj *fala* (forms confusable with *falar* 'talk')

This, too, cannot be the whole story, as the set of 40 defective verbs compiled in the Aurélio dictionary and Cunha & Cintra's grammars contain many which do not submit to this explanation. Perhaps, indeed, we are relegated to a position no more intriguing than that of Maiden & O'Neill (2010), who state "The major domain of defectiveness (present subjunctive and 1sg present indicative) seems, and is, irreducibly arbitrary if one seeks a motivation outside the morphological system itself."

However, a phonological generalization has recently been proposed, at least for Portuguese. Postma (2013) noticed that of the 40 defective verbs listed in Aurélio and Cunha & Cintra, 35 (87%) have a coronal sonorant (n,l,r,ɲ) immediately following the stem-final vowel (e.g. *colorir, banir*, *polir*).

In terms of the explanation Postma offers (which we consider to be independent of the empirical generalization itself, and henceforth focus only on the latter in the present paper), his idea is that deleted theme-vowel harmony (e.g. *dormir* (inf.)*, dormes* (2sg. ind) vs. *durmo* (1sg ind)*, durmas* (2sg sbj); Harris 1974) competes with the coronal for positioning within the root. As Harris (1974) proposes that the vowel raising found in these forms is the result of a kind of autosegmental spreading of the deleted theme vowel to the rhizotonic position, Postma suggests that coronal sonorants positioned immediately after the would-be rhizotonic vowel block and compete the docking of the theme vowel, and that the result is a grammatical crash, leading to ineffability. While we must refer the reader to Postma's work for the details of the proposal, the intuition is that the fact that defective verbs are defective precisely where the theme vowel /i/ cannot be overtly realized has an interaction with the presence of a coronal sonorant following the final root-vowel.

The question we therefore wished to examine was whether indeed, speakers internalize such a generalization. In fact, before doing such wug-experiments, we need to know what the actual list of defective verbs *is* for

real speakers. For example, are *explodir* 'explode' and *latir* 'bark' really defective (and hence exceptions to Postma's generalization)? In order to conduct this study, we took all 40 defective verbs, removed those that two PhD students did not know the meanings of. This left 21 verbs. We asked participants to rate familiarity of each verb from 1-5 in a pretest and excluded those with a mean lower than 2.5. Participants (n=24) produced the finite 1pl.ind and the 3sg subjunctive, and then rated their confidence in their own productions, following the procedure of Albright (2009), who used the measure of a speaker's confidence in his/her own production as a measure of defectiveness. Importantly, our study (unlike others, including Albright's) specifically focused on the defectiveness of 3sg subjunctive, in order to see how truly general the pattern was, and avoid the potential confound of 1sg semantic restrictions. Details that cannot be included within scope of the present chapter are to be found in Nevins, Damulakis & Freitas (in preparation), to which the reader is referred for a more extended report and discussion.

The results are shown in Figure 11, where a statistical comparison of the mean rating for the 1pl ind vs. the 3sg sbj of each verb is submitted to a t-test. Forms in red are thus defective. As the results show, *latir* and *explodir* were not defective, i.e. not significantly different in their inflected 1pl ind vs. 3sg sbj forms. By contrast, all of the forms that were defective have a coronal sonorant in the root, upholding Postma's generalization. Importantly, these defective verbs can have a mid vowel, a high vowel, or a low vowel in the root, thereby rendering the generalization about the nature of the consonant more predictive than about, say, having mid-vowels (an explanation often proposed in the literature). Now that we have found *which* the set of defective verbs in Portuguese actually are, in future work, we can examine whether the phonologically-based generalization is extended to novel verbs or not.

6. Conclusion

This paper has covered a range of phenomena within the patterns of syncretism, defectivity, and allomorphy found within Portuguese derivational and inflectional morphology. While focusing on productivity, we have highlighted not only on the productivity of certain morphological formations, but on a more remote question: the productivity of certain generalizations by learners. We have found that certain generalizations that seem arbitrary and unnatural, such as the L-shaped pattern, are not productive, due to asymmetric growth in the verbal vocabulary size. We have found that conjugation-specific knowledge holds, when it has a principled basis.

Moreover, we have found that featural decomposition (i.e. identity of form within the 2nd person across indicative and subjunctive moods), Markedness-Neutralization (i.e. neutralization of the theme vowel in the 2nd

and 3rd conjugation when deverbalized), isomorphism between structure and spellout (i.e. favoring the short-form participles when the argument structure of the verb is comparatively reduced) and phonologically-based regularities among defective verbs are principled, interface-grounded properties that play a guiding hand in morphological naturalness, as assessed by controlled experimental comparisons. Naturally, the interplay between theoretically-driven hypothesis formation based on the interface demands on morphology and the use of experimental methods such as elicited production, forced-choice rating, and confidence ratings on production represent possibilities for research into language structure that are by no means limited to Portuguese nor the Romance languages.

References

Albright, Adam. 2009. Lexical and morphological conditioning of paradigm gaps. In *Modeling ungrammaticality*, ed. Curt Rice and Sylvia Blaho, 117–164. London: Equinox Publishing.

Anderson, Stephen. 1981. Why Phonology Isn't 'Natural'. *Linguistic Inquiry* 12: 493-539.

Anderson, Stephen. 1992. *A-morphous Morphology*. Cambridge University Press.

Anshen, Frank, and Aronoff, Mark. 1997. Morphology in real time. *Yearbook of Morphology* 9–12.

Aronoff, Mark. 1994. *Morphology by Itself.* Cambridge, MA: MIT Press.

Bachrach, Asaf & Andrew Nevins (eds). 2008. *Inflectional Identity*. Oxford University Press.

Becker, Michael, Nihan Ketrez, and Andrew Nevins. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. *Language* 87:84–125.

Berent, Iris, Tracy Lennertz, Paul Smolensky, and Vered Vaknin-Nusbaum. Listeners' knowledge of phonological universals: evidence from nasal clusters. *Phonology* 26: 75-108.

Blevins, Juliette. 2004. *Evolutionary Phonology*. Cambridge University Press.

Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology*. Cambridge, MA: MIT Press.

Bourne, L.E. 1970. Knowing and using concepts. *Psychological Review* 77: 546-556.

Chagas, Paulo. 2007. Athematic Participles in Brazilian Portuguese: A Syncretism in the Making. *Acta Linguistica Hungarica* 54: 119–128.

Dressler, Wolfgang, Willi Mayerthaler, Oswald Panagl & Wolfgang Wurzel. 1987. *Leitmotifs in Natural Morphology.* Amsterdam: John Benjamins.

Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35: 355-292.

Feldman, Jacob. 2000. Minimization of Boolean complexity in human concept learning. *Nature* 407: 630-633

Finley, Sara. 2012. Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes* 27: 1550-1562.

Freitas, Maria Luisa. 2014. Two nominalizing suffixes in Brazilian Portuguese: locality constraints on morphophonological realization. Generals paper, UNICAMP.

Freitas, Maria Luisa & Nevins, Andrew (in prep). Conjugation-specific knowledge and nominalizing affix choice. Unpublished manuscript.

Graf, Thomas. 2012. An Algebraic Perspective on the Person Case Constraint. *UCLA Working Papers* 17: 85-90.

Hale, Mark & Reiss, Charles. 2008. *The Phonological Enterprise.* Oxford University Press.

Harris, James. 1974. Evidence from Portuguese for the "Elsewhere Condition" in Phonology. *Linguistic Inquiry* 5: 61-80.

Hayes, Bruce, Donca Steriade & Robert Kirchner (eds.). *Phonetically-Based Phonology*. Cambridge University Press.

Heinz, Jeffrey. 2010. Learning Long-Distance Phonotactics. *Linguistic Inquiry*, 41:623-661.

Koontz-Garboden, Andrew. 2008. Monotonicity at the lexical semantics-morphosyntax interface. In Emily Elfner and Martin Walkow (eds.), *Proceedings of the 37th annual meeting of the Northeast Linguistic Society, Volume 2*. Amherst, MA: GLSA. pp. 15–28

Lindsay, Mark and Mark Aronoff. 2013. Natural Selection in Self-Organizing Morphological Systems. In *Morphology in Toulouse: Selected Proceedings of Décembrettes 7*, ed. by Fabio Montermini, Gilles Boyé, and Jesse Tseng. Germany: Lincom Europa.

Lobato, Lucia. 1999. Sobre a Forma do Particípio do Português e o Estatuto dos Traços Formais. *DELTA* 15.1

Maiden, Martin. 2005. Morphological autonomy and diachrony. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology* 137–175.

Maiden, Martin & Paul O'Neill. 2010. On Morphomic defectiveness: Evidence from the Romance languages of the Iberian Peninsula. In *Defective Paradigms: Missing forms and what they tell us*, ed. by Matthew Baerman, Greville Corbett, and Dunstan Brown. Oxford University Press, pp. 103-124.

Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25:83–127.

Nevins, Andrew, Damulakis, Gean & Freitas, Maria Luisa (in prep). Phonological regularities among defective verbs. Unpublished manuscript.

Nevins, Andrew & Rodrigues, Cilene. 2012. Naturalness biases, 'morphomes', and the Romance First Person Singular. Paper presented at IGG, Sevilla.

Nevins, Andrew & Rodrigues, Cilene. 2014. "The L-syntax of athematic past participles : evidence from wug experiments." Submitted to *Revista da ABRALIN.*

Oltra-Massuet, Isabel. 1999. On the notion of theme vowel: A new approach to Catalan verbal morphology. MIT Occasional Papers in Linguistics 19. Cambridge, Mass.: MITWPL.

Pertsova, Katya. 2011. Grounding systematic syncretism in learning. *Linguistic Inquiry* 42: 225-266.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge, MA: MIT Press.

Postma, Gertjan. 2013. Metaphonic Blocking in Portuguese as a Linearization Deadlock. Workshop on Metaphony, Meertens Institute.

Tang, Kevin & Andrew Nevins. 2013.Quantifying the Diachronic Productivity of Irregular Verbal Patterns in Romance. *UCL Working Papers in Linguistics* 25.

Steriade, Donca. (2008) A pseudo-cyclic effect in Romanian morphophonology. In Asaf Bachrach and Andrew Nevins (eds), *Inflectional identity*. Oxford University Press, 313-359.

Yang, C. (2005). On productivity. *Linguistic Variation Yearbook, 5(1),* 265–302.

Zimmer, Karl. 1969. Psychological correlates of some Turkish Morpheme Structure Conditions. *Language* 45.2:309–321.

Figure 1: Results of L-Morphome Choice for Portuguese (n=250), given 15 stimuli over which to make the choice

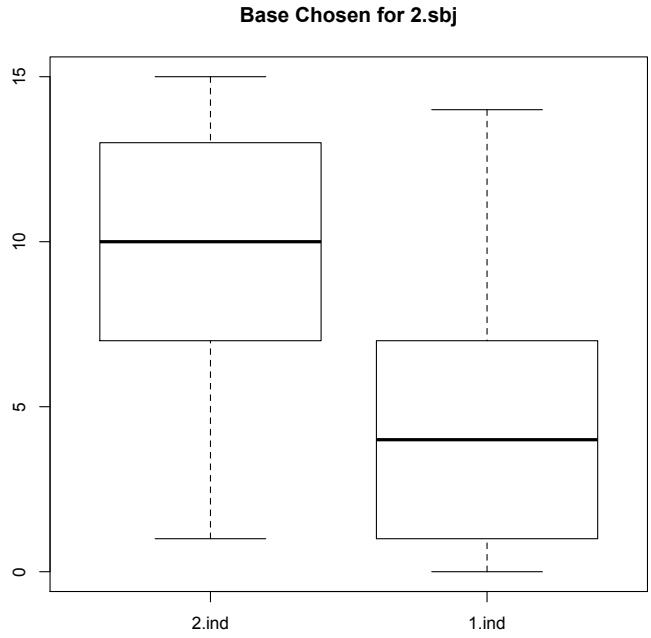

Figure 2: Results of L-Morphome Choice for Italian (n=175), given 15 stimuli over which to make the choice

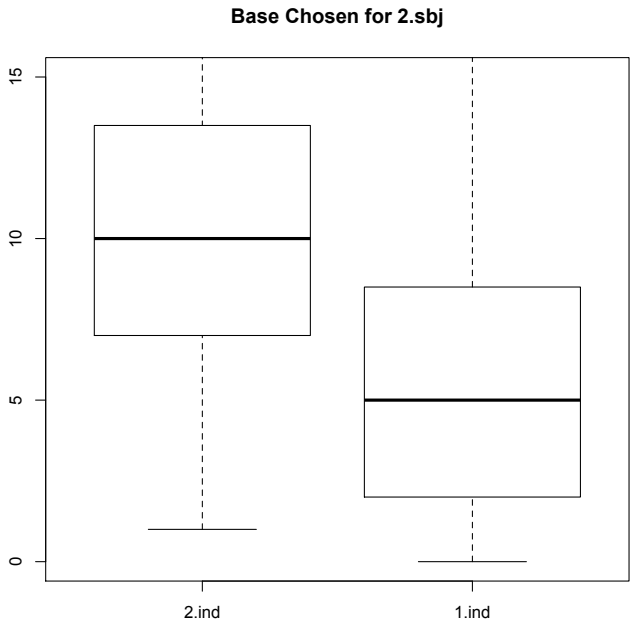

Figure 3: Results of L-Morphome Choice for Spanish (n=148), given 15 stimuli over which to make the choice

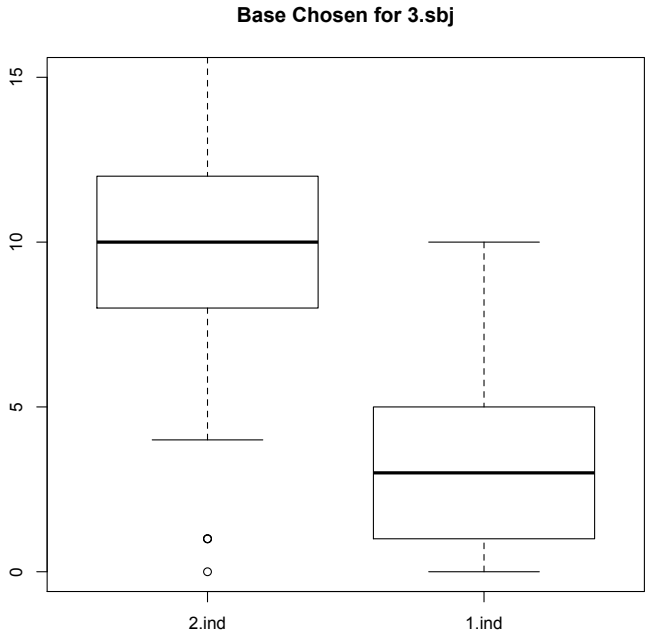

Figure 4: Growth of ratio of 1st conjugation verbs in Portuguese, 1525-1899 (*Colonia* corpus, 5.1 billion words)

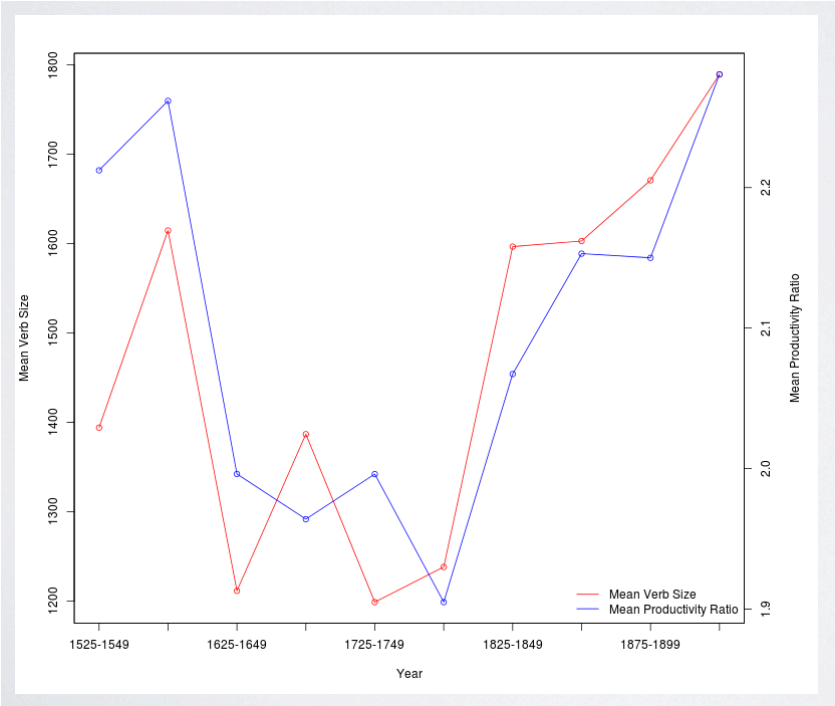

Figure 5: Growth of ratio of 1st conjugation verbs in Italian, 1550-1974 (Google N-Gram, 40 billion words)

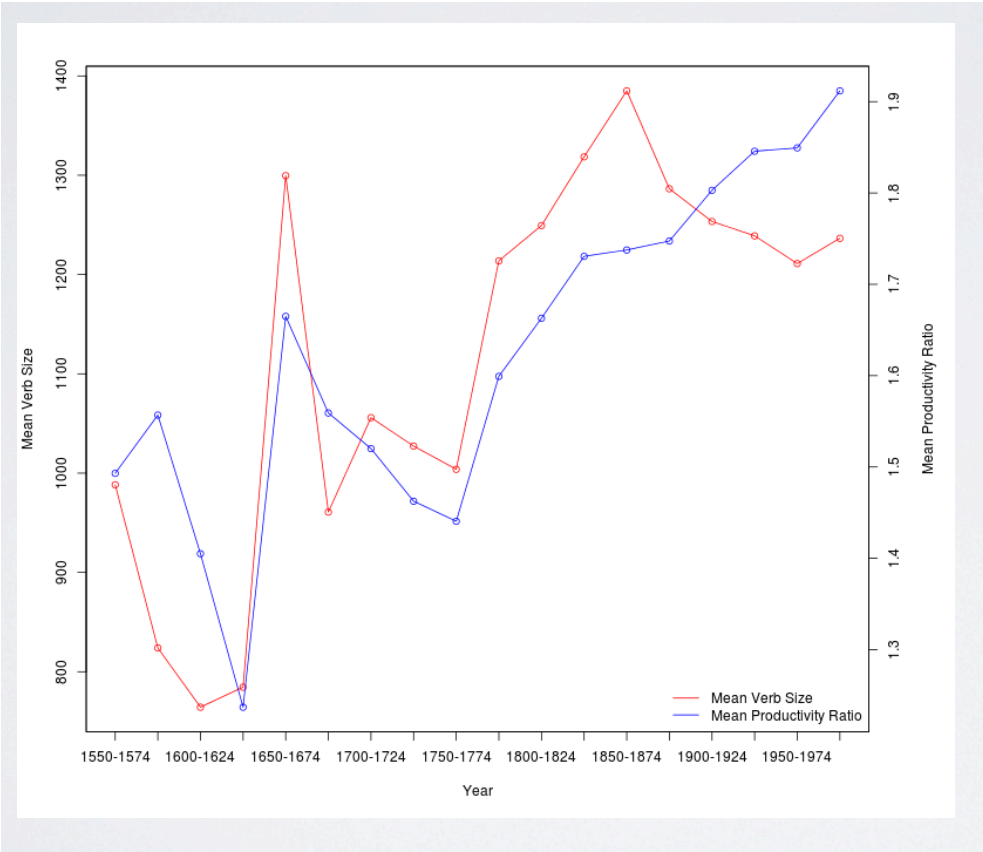

Figure 6: Growth of ratio of 1st conjugation verbs in Spanish, 1522-1996 (Google N-Gram, 83 billion words)

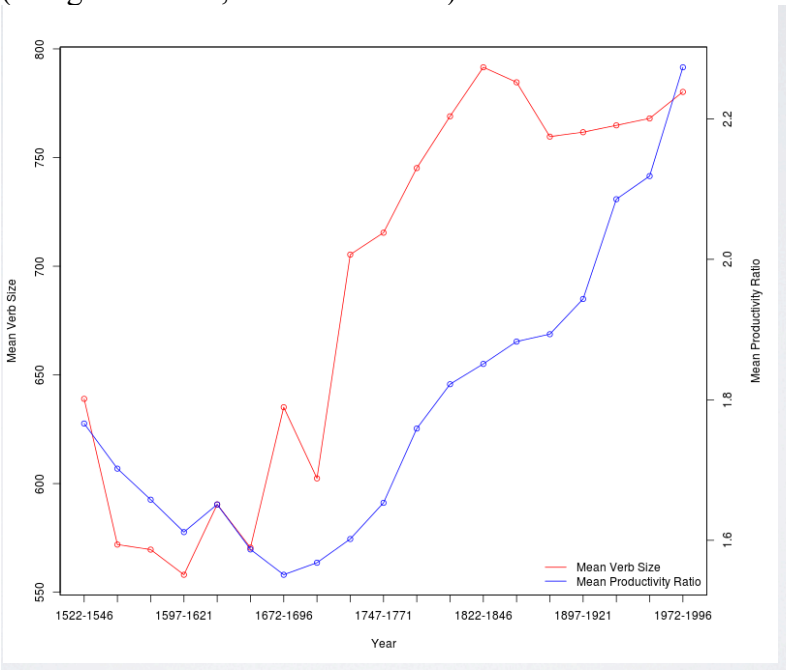

Figure 7: Preferences for *-ção* and *-mento* in the Houaiss dictionary (2,085 forms)

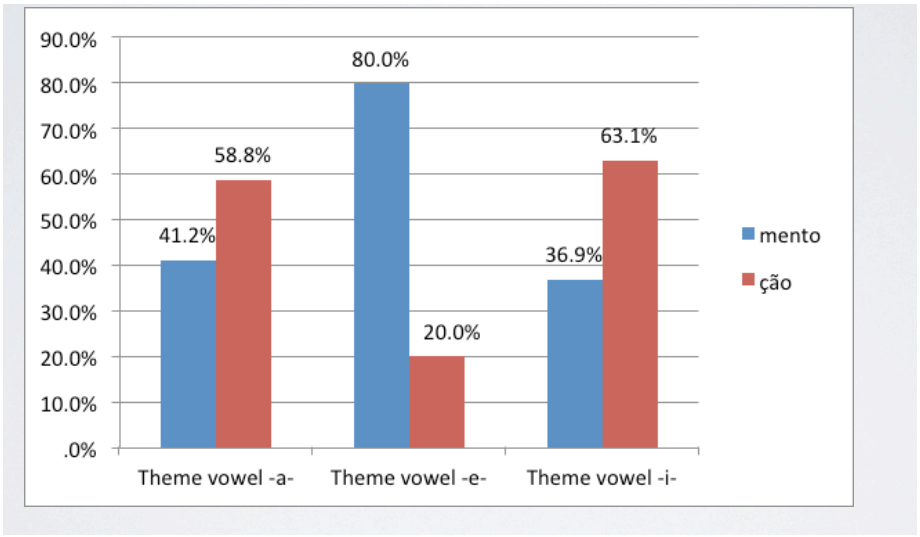

Figure 8: Mean rating preference for *–mento* forms (scale of 1-5), when it raises theme vowel (n=140)

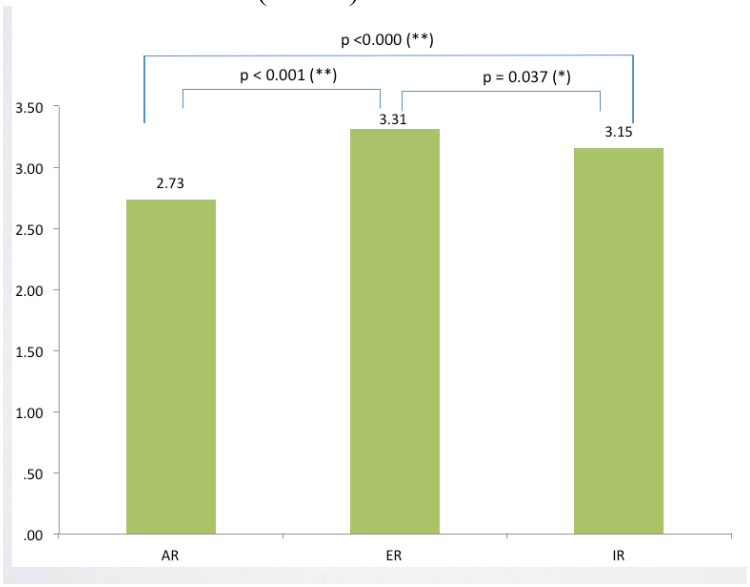

Figure 9: Mean rating preference for *–mento* forms (scale of 1-5), when it doesn't raise theme vowel (n=103)

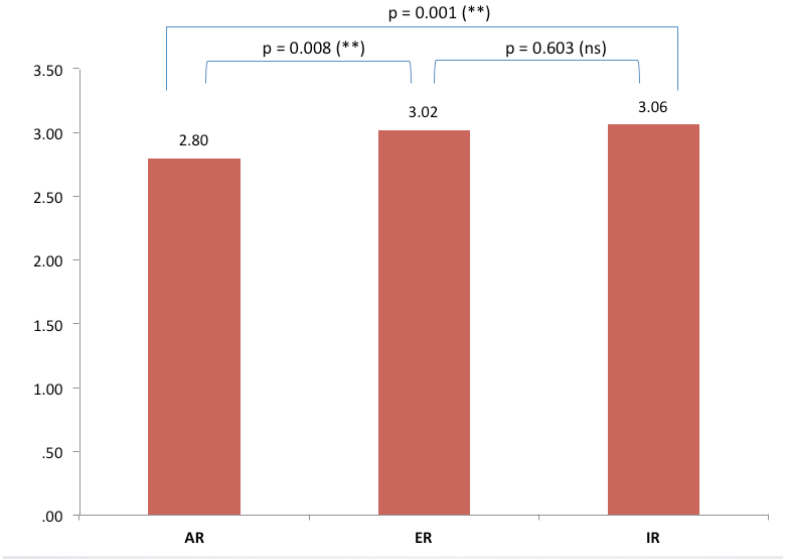

Figure 10: Mean rating preference for short-form participles (scale of 1-5), by verb class (n=98)

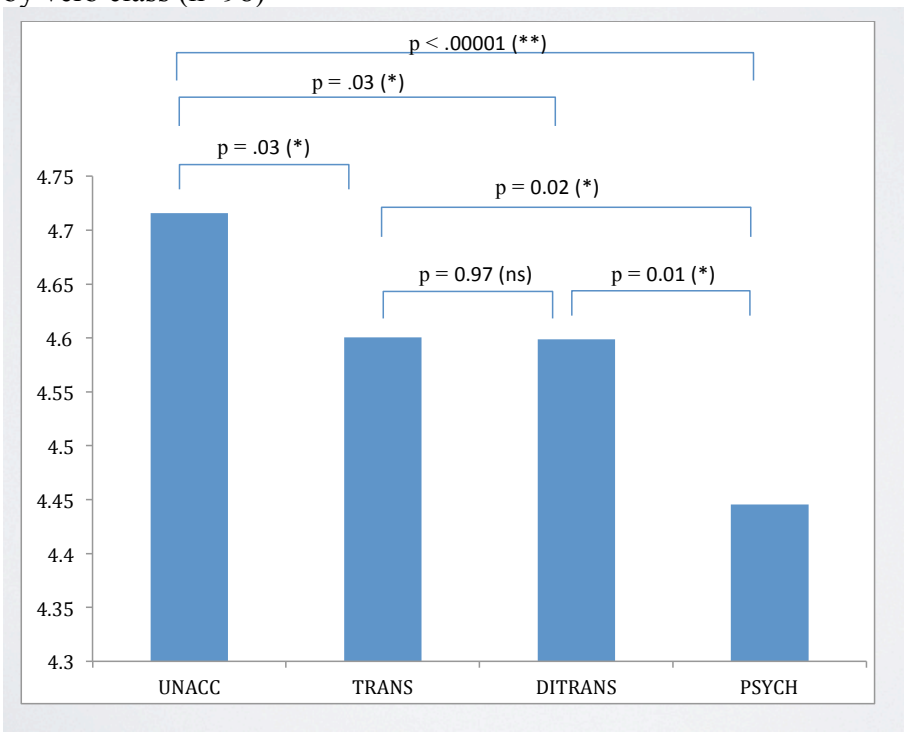

Figure 11: Mean rating preference for finite forms of putative defective verbs, scale of 1-5 (n=24)

Verb		3Sbj	1Pl	p-value
ruir	collapse	1.39	2.90	.000
grunhir	grunt	1.46	1.96	.073
polir	polish	1.48	3.26	.000
parir	give birth	1.65	3.50	.000
falir	fail	1.95	3.73	.000
banir	banish	1.95	3.57	0.000
imergir	immerse	2.00	2.86	0.029
abolir	abolish	2.00	3.46	0.000
demolir	demolish	2.00	3.75	.001
exaurir	exhaust	2.13	2.46	.277
colorir	color	2.33	4.11	0.000
florir	blossom	2.44	3.24	.078
emergir	emerge	2.64	2.89	0.534
ungir	anoint	2.74	3.36	.154
latir	bark	3.00	3.65	.103
engolir	swallow	4.08	4.41	.182
explodir	explode	4.33	4.44	.682