



Open Library of Humanities

Implementing the U-shaped curve in Distributed Morphology and as a by-product of the third factor

Paulo Ângelo Araújo-Adriano, Universidade Estadual de Campinas (UNICAMP), Brazil, pauloangeloaa@gmail.com

Rafael Luis Beraldo, Universidade Estadual de Campinas (UNICAMP), Brazil, rberaldo@cabaladada.org

In this paper, we present a novel formal approach to overregularisation (e.g. *fazi* “doed”, *trazi* “brought”) in Brazilian Portuguese (BP), a phenomenon widely attested in children’s production. Three previous accounts for this phenomenon are also reviewed and at least partially rejected. The first explained overregularisation on the basis of Minimize Exponence (an economy principle in Distributed Morphology) not being followed by children; the second achieved the same goal by positing constraint ranking (a main feature in Optimality Theory) that is faulty in children’s grammar when compared to adult knowledge; and the third analyses BP through the framework of Grammar Competition. Our main departures with these approaches are that the first two a) are dependent on children’s breaking of the theoretical principles just mentioned; b) no learning mechanism is ever discussed to justify this; c) although both describe a system that indeed produces overregularised forms, they fail to predict or capture the three stages of the U-shaped curve, being an explanation solely of its dipping section and not of children’s development; and that the third d) only deals with noise but not exceptions, the main point of the overregularisation process. In an attempt to remedy those perceived flaws, our approach keeps theoretical principles intact, maintaining parity with adult competence and addressing points a) and b). Furthermore, we integrate the chosen descriptive toolkit of Distributed Morphology and general cognitive biases/abilities in the form of Biberauer’s Minimise Maximise Means and Yang’s Tolerance Principle (TP) in order to explain the developmental path of BP acquisition takes throughout the U-curve, thus addressing points c) and d). Finally, we argue that our proposal is naturally compatible with the notion of grammar competition by virtue of assuming TP, which might explain why children do not categorically overregularise irregular verbs, but rather do so at a given rate.



1. Introduction

That children overregularise irregular verbs is a matter of fact in the language acquisition process. Between 3 and 5 years of age, every child produces irregular verbal forms, such as *trazi* ‘bringed’ (for *trazer* ‘bring’) as if it were a regular verb, such as *vivi* ‘lived’ (for *viver* ‘live’) and *comi* ‘lived’ (for *comer* ‘ate’) in the case of Portuguese. This developmental process is known as the U-shaped curve (Pinker; Prince, 1988) because it occurs in three stages: in the first, the infant produces irregular verbs according to the adult’s grammar – *trouxe* ‘brought’; in the second stage, surprisingly, it overgeneralises the rule for the regulars and applies it to the irregulars, even though this is not a categorical process, and, finally, in the third stage, the child returns to producing the target grammar verbal forms. This process is illustrated in (1).

- (1) 1st stage: *trouxe* ‘brought’, *soube* ‘knew’, *comi* ‘ate’, *vivi* ‘lived’
 2nd stage: *trazi* ‘bringed’, *sabi* ‘knewed’, *comi* ‘eated’, *vivi* ‘lived’
 3rd stage: *trouxe* ‘brought’, *soube* ‘knew’, *comi* ‘ate’, *vivi* ‘lived’

The child’s recognition of a particular pattern of regular verb constitution, and its application to the irregular verbs, is a striking clue that the child, from an early age, notices and manipulates the pieces – morphemes – that compose a word. That is, the errors they make end up being a signal of their language analysis. Works dedicated to investigating this topic consider that what the child does is nothing more than extending the verbal structure of regulars, applying it to irregulars. Descriptively speaking, following Camara Jr.’s (1970) verbal structure for Brazilian Portuguese, children’s overregularised verbs follow the same template as regular verbs (cf. section 2). The mismatch between child and adult grammars would result from the fact that languages present exceptions to productive rules, that is, from factors idiosyncratic to each language.

The issue this paper tries to answer may be boiled down to a single question: can the formal toolkit available for describing language explain such behaviour? Within a formal perspective, the three proposals available for (Brazilian) Portuguese do not account for the U-shaped curve in its totality. Takahira’s (2013) assumes Distributed Morphology (DM), proposing that the child produces overregularised forms because the verb does not move to T; thus, Minimize Exponence (Siddiqi, 2009), the principle/constraint of forming more complex heads in order to receive less morphological content, is not followed. On the other hand, Lorandi’s (2007, 2010) proposal, based on the Optimality Theory, explains the errors as a restriction of fidelity to the verbal root, ranked in a higher position than the non-fidelity to the verbal root. The author assumes that the verbal root base for the lexeme *fazer* ‘to do’ and *trazer* ‘to bring’ is *faz* and *traz*, respectively (p. 154). Hence, following the root fidelity restriction, the child would produce overregularised forms that are faithful to the base, such as *fazi/trazi*, differently from the adult’s grammar, whose higher restriction that generates the grammatical form and therefore the best candidate among the ranking is the non-fidelity restriction (*fiz/trouxe* from *faz/traz*).

A third work dealing with the issue of regular and irregular verb acquisition in Brazilian Portuguese is Wuerges' (2014). In her thesis, she analysed the production of children aged 1;6 to 4;0 in order to verify the role of input frequency. All verb forms present in the input to a specific child were described in terms of the rules producing each verb tense. In particular, 41 different verb classes were identified as composing the possible irregular verb inflections. Wuerges then applied Yang's Variational Model (2002) to successfully predict which irregular verbs are more prone to overregularisation given their frequency in the input.

The main problem with those proposals is that they do not entirely explain the U-shaped curve. The first two account for the period when the child makes errors (2nd Stage), but not for what happens after (3rd Stage) and, most notably, before that (1st Stage). Additionally, they make unobserved predictions. For instance, the moving of *v* to T in child language was attested by Santos and Lopes (2017) in a period that includes 2nd Stage, which would wreck Takahira's (2013) proposal. Similarly, Lorandi's (2010) proposal of different rankings in child and adult grammar does not explain the early stage of target-form production, nor does it provide a mechanism for explaining *what* triggers the subsequent rank alternations leading to the stable, final grammar. Finally, Wuerges' (2014) work follows Yang's (2002) Variational Model of Language Acquisition, which accounts only for noise but not for exceptions. This flaw is even commented on and observed by Yang (2016) himself, who claims that this model is inappropriate in this respect (cf. discussion in Section 2).

Given this explanatory incompleteness, this article proposes an alternative theoretical analysis for the U-curve as a whole. As Takahira (2013), we will adopt the apparatus of Distributed Morphology, which considers that what is typically called a "word" is, in reality, a product of syntactic, semantic, and phonological operations. In addition, we will use a formal model of language acquisition, Yang's (2016) Tolerance Principle, which states that the child's Language Acquisition Device (LAD) is sensitive to the statistics in the data that nourish their Universal Grammar (UG). In other words, children would generalise productive rules and list exceptions depending on the amount of input they are exposed to. In a nutshell, we can account for the U-shaped curve of acquisition of irregular verbs as a by-product of the third factor (Chomsky, 2005) in language design: a matter of acquisition strategy that leads the child to follow the same rule, which ends up accidentally being an undesired outcome.

We argue, against Takahira, that the child follows the same rule set as adults do. However, rule productivity depends on the amount of input to which children have been exposed. This leads to overgeneralisations or the correct listing of exceptions. Thus, in having accumulated a small amount of data, the child lists all the verbs in the perfect past tense (*comi* 'ate', *bebi* 'drank', *fiz* 'did', *comprei* 'brought' etc.), because, as per the Tolerance Principle, the relative difference between (ir)regulars do not license the postulation of a rule. Therefore, the correct production of irregulars is predicted. Having accumulated more data (around two years of age – Ferrari-Neto & Lima, 2015) and,

consequently, with a higher proportion of regulars over irregulars, the child generalises an overly productive rule, producing irregulars **fazi*, **trazi* as the regulars *comi*, *bebi*. Finally, approaching adult grammar, the proportion of irregulars over regulars diminishes. The child starts to perceive exceptions to the general rule, leading them to list such exceptions next to the prolific rule.

Here we use “listing” in the sense of Yang (2016) *and* in the sense of DM. According to the former, if a postulated rule has too many exceptions, it is cognitively more efficient to just “lexically list everything” (p. 41). However, the actual architecture of this list, one could argue, is heavily theory-dependent. As one of DM’s main features is formally describing the nature of the lexicon, which is viewed as a set of three separate lists, this theoretical model seems naturally suitable for actually implementing a representation of the mechanism put forth by Yang. Under this view, “listing” should not be understood as a running record of verbs in the lexicon, but as a listing of pairs of features and phonological exponents. Thus, our work is, in a way, an exploration of the theoretical casting of the term “listing”: the task in language acquisition is to pair features with phonological material (for theoretical and empirical consequences on whether features are specified or not by Universal Grammar, cf. Roberts, 2019, pp. 99–101). Conversely, by using no more than the basic tools from DM and showing they can adequately capture the U-curve, we are putting this model to the test. There has been a *desideratum* (Chomsky, 1966, 1986; Yang & Roeper, 2011) that a theory of language must precisely explain how acquisition occurs. Similarly, it has to be compatible not only with the “synchronic” adult knowledge of language, but also with the “diachronic” states that children’s intermediary grammars may take up. In this manner, this work is also about putting DM through its paces.

Anchored in the principles above, namely Yang’s Tolerance Principle and in DM, we (i) can predict when the general rule becomes productive and when the exceptions are listed, (ii) can propose that overregularisation is triggered by the same rule as regular verbs, returning the property of discrete infinity to the child; and, finally, (iii) explain all stages of the U-curve. These points are made explicit as follows. Section 2 presents previous observations of the process of acquiring irregular verbs, in addition to theoretical-formal proposals for the phenomenon. Then, in section 3, the models of Distributed Morphology and the Tolerance Principle are discussed. Next, the alternative proposal is made in section 4, and, finally, in the last section, we make our final remarks.

2. Some facts and proposals about irregular verb acquisition

It is commonly assumed that children are able to morphologically analyse input when errors can be observed in their productions, as proposed by Bowerman (1982). Regarding irregular *versus* regular productivity, the topic of this paper, a typical regular Portuguese verb is derived by the suffixation of tense/mood and number/person morphemes into the verb root. Camara Jr. (1970) defines the verbal morphologic constitution as in (2), where T from Theme is the product of Root and a Theme Vowel (TV) added to some Inflectional Suffix, such as the gathering of the Tense-

Mode suffix (TM) and Person-Number suffix (PN). For any regular verb, prototypical TV could be divided according to its conjugation endings: the first, second and third conjugation end, respectively, in *-a* (amar ‘love’), *-e* (beber ‘drink’) and *-i* (partir ‘leave’).

(2) Regular verb

T (R + TV) + IS (TM + PN)

A simple regular verb in the present tense in Brazilian Portuguese (BP) has a zero morpheme \emptyset for TM and an *-o* PN. Due to a phonological restriction, all unstressed vowels in the final position are dropped in the presence of another vowel, thus to derive *bebo* ‘I drink’, the first person singular in the simple present tense, one adds to $\sqrt{\text{BEB}}$ the /e/ thematic vowel (from the first *-e* conjugation ending), the \emptyset TM morpheme and the *-o* PN. With the /e/ reduction in the presence of another vowel, namely /o/, the verb is realised as *bebo*.

(3) R + TV + TM + PN

$\sqrt{\text{BEB}}$ + *-e* + \emptyset + *-o* = *bebo*

One of the common errors that children make is the overregularisation of irregular verbs, which exposes very young children’s knowledge of the difference between a root and an affix: Lorandi (2010) shows that for every error in the acquisition of Brazilian Portuguese, the root (which we denote by “ $\sqrt{\text{V}}$ ” from now on) was maintained, with the addition of a fitting inflection affix. The examples in **Table 1** below from the author suggest that children know the verb pieces (morphemes) because $\sqrt{\text{FAZ}}$ was preserved in all occurrences, and a correct inflection affix was used.

| Structure produced | Adult grammar | Regular verbs |
|---|---|--|
| faz-i (2:6) <i>do/make</i> . _{IND.PST.1.SG} ‘doed/maked’ | fiz <i>do/make</i> . _{IND.PST.1.SG} ‘did/made’ | varr-i <i>sweep</i> . _{IND.PST.1.SG} ‘swept’ |
| faz-o (3:6) <i>do/make</i> . _{IND.PRS.1.SG} ‘do/make’ | faço <i>do/make</i> . _{IND.PRS.1.SG} ‘do/make’ | varr-o <i>sweep</i> . _{IND.PRS.1.SG} ‘sweep’ |
| faz-esse (3:11) <i>do/make</i> . _{SUB.PST.1.SG} ‘doed/maked’ | fizesse <i>do/make</i> . _{SUB.PST.1.SG} ‘did/made’ | varr-esse <i>sweep</i> . _{SUB.PST.1.SG} ‘swept’ |
| faz-eu (4:0) <i>do/make</i> . _{IND.PST.3.SG} ‘doed/maked’ | fez <i>do/make</i> . _{IND.PST.3.SG} ‘did/made’ | varr-eu <i>sweep</i> . _{IND.PST.3.SG} ‘swept’ |

Table 1: Overregularisations attested in children’s productions, target form, and template used for overgeneralisation. Data in the first column come from Lorandi (2010).

Another piece of evidence for the “child morphologist” is their production of a non-existent verb from a typical nominal root, such as *borrachar* ‘to rubber’ – to erase with a rubber –, *xizar* ‘to X’ – to tick with an X –, *vassourar* ‘to broom’ – ‘sweep’ – (cf. Lorandi, 2010, p. 87), which shows a sensibility for the morphological structure. Numerous works, including with an empirical focus (cf. Ferrari-Neto & Lima, 2015; Figueira, 1996, 2010; Lorandi, 2010; Maldonade, 2003; Takahira, 2013) have attested that children produce the irregular *saber* (‘know’) as the overregularised *sabo* (‘know’) instead of the target irregular form *sei* (‘know’) for the present indicative tense. Applying the descriptive verb-making template just described to *saber*, we see that *sabo* clearly fits into the regular verb class in (3): in this sense, the overregularisation observed for the present tense could be explained by the fact that the child learnt the pattern in (3) for the simple present tense formation and applies it to the verbs they face.

- (4) (a) 2;08.14 (Maldonade, 2003, p. 146)¹
 Adult: Será que eu sei, Marcela, faze uma casa de massinha? Vamo vê!
 ‘I wonder if I know, Marcela, to make a house out of play dough. Let’s see!’
 Child: Eu **sabo**.
 ‘I know.’
 Adult: É, você é danada! Que cor você qué?
 ‘You are so smart! What colour do you want?’
 Child: Azul.
 ‘Blue.’
- (b) R + TV + TMS + NPS
 $\sqrt{\text{SAB}} + -e + -\emptyset + -o = \text{sabo}$ (cf. (3))

As for the simple past tense, the structure of the regular verb would also follow the same structure in (3), but in this case, the TV for a second conjugation verb such as *comer* is *-i*, and the TMS is again zero $-\emptyset$. In the first person singular of past simple, *-i* is neutralised between the 2nd and the 3rd conjugation. Since the NPS is /i/, both vowels will be fused (Camara Jr., 1970, p. 159). Following this rule, a child also makes “mistakes”, reapplying it to irregular verbs such as *trazer* ‘to bring’. Then, it is plausible to assume that, by producing *trazi*, the child is analysing this verb as regular since the rule for the regulars is being strictly followed.

- (5) (a) R + TV + TMS + NPS
 $\sqrt{\text{COM}} + -i + -\emptyset + -i = \text{comi}$
- (b) R + TV + TMS + NPS
 $\sqrt{\text{TRAZ}} + -i + -\emptyset + -i = \text{trazi}$

¹ In all examples of conversations with children within this paper, we did not make any interventions regarding the methodology used to transcribe the data. We kept the authors’ original format.

Other examples, such as *fazi* ‘doed’, *sabo* ‘knowed’, *queri* ‘wanted’, *cabeu* ‘fitted’, *trazeu* ‘bringed’, follow the same structure, as if the child analysed irregular verbs as regulars. This overregularisation strongly suggests that children notice that a verb is a compound of different pieces (the morphemes), and that is not an atomic unity, a process found in different languages. In English, for instance, children generalise the *-ed* rule, applying it to irregular *teach*, producing **teached* instead of *taught* (cf. Marcus et al., 1992). However, the compelling aspect about this (universal) process is that it occurs in a sequence of hits, misses, and later more hits, showing a non-linear development; hence this learning curve became known as the U-shaped curve, illustrated in **Figure 1**.

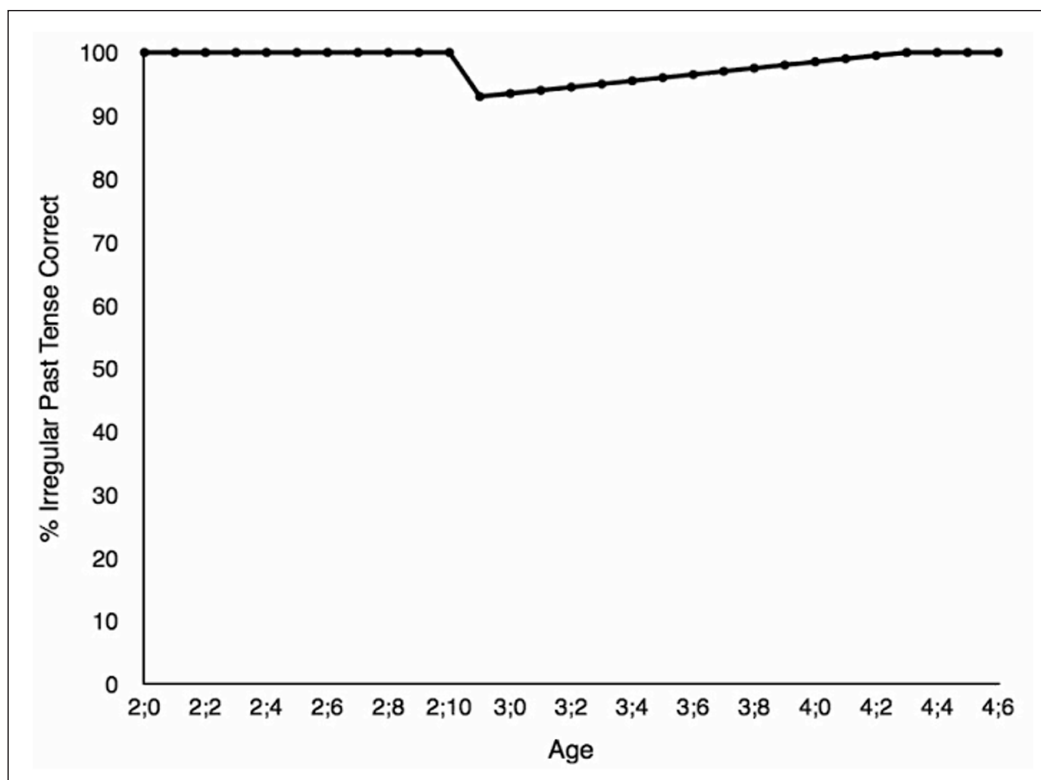


Figure 1: U-shaped curve of past-tense development in one child's linguistic development. Adapted by Yang (2016, p. 88) from Marcus et al. (1992).

Having presented these facts about the acquisition of (ir)regular verbs, we now feature three proposals attempting to explain this process. Lorandi (2007, 2010), for example, assumes Optimality Theory – OT – (Prince & Smolensky, 1993), according to which grammars are explained by conflicting restrictions, ranked descending from left to right in the *tableau* and subject to violations. Different from the mainstream Generative Grammar point of view, under OT, Universal Grammar is composed of a set of constraints that can be violated (Prince & Smolensky

1993, p. 3). Among all possible candidates, the one violating the fewest restrictions succeeds. Regarding overregularisation, Lorandi (2007, 2010) suggests that one of these restrictions is related to the fidelity of root (OOROOTFAITH): the constraint of fidelity of root is ranked in a higher position (closer to candidates on the *tableaux* in **Tables 2** and **3** below). Thus, for *trazer* ‘bring’ and *fazer* ‘do/make’, whose verbal root bases are, respectively, *traz* and *faz*, as mentioned before, if OOROOTFAITH is violated – represented by * – the output is the adult’s grammar *trouxe* ‘I brought’ and *fiz* ‘I did/made’ (the pair in b) below), since the root changed from $\sqrt{\text{TRAZ}}/\sqrt{\text{FAZ}}$ to $\sqrt{\text{FIZ}}/\sqrt{\text{TROUX}}$, but if the restriction OOROOTFAITH is preserved, then the best candidate is *trazi* and *fazi* (traditionally represented by ☞ in the *tableaux*). In this case, given that the adult’s grammar form violates a higher-ranked restriction, it is eliminated, represented by !. The following *tableaux* illustrate that:

| Base | Candidate | OOROOTFAITH | ~OOROOTFAITH |
|---------|-----------|-------------|--------------|
| a) traz | ☞ trazi | | * |
| b) traz | trouxe | *! | |

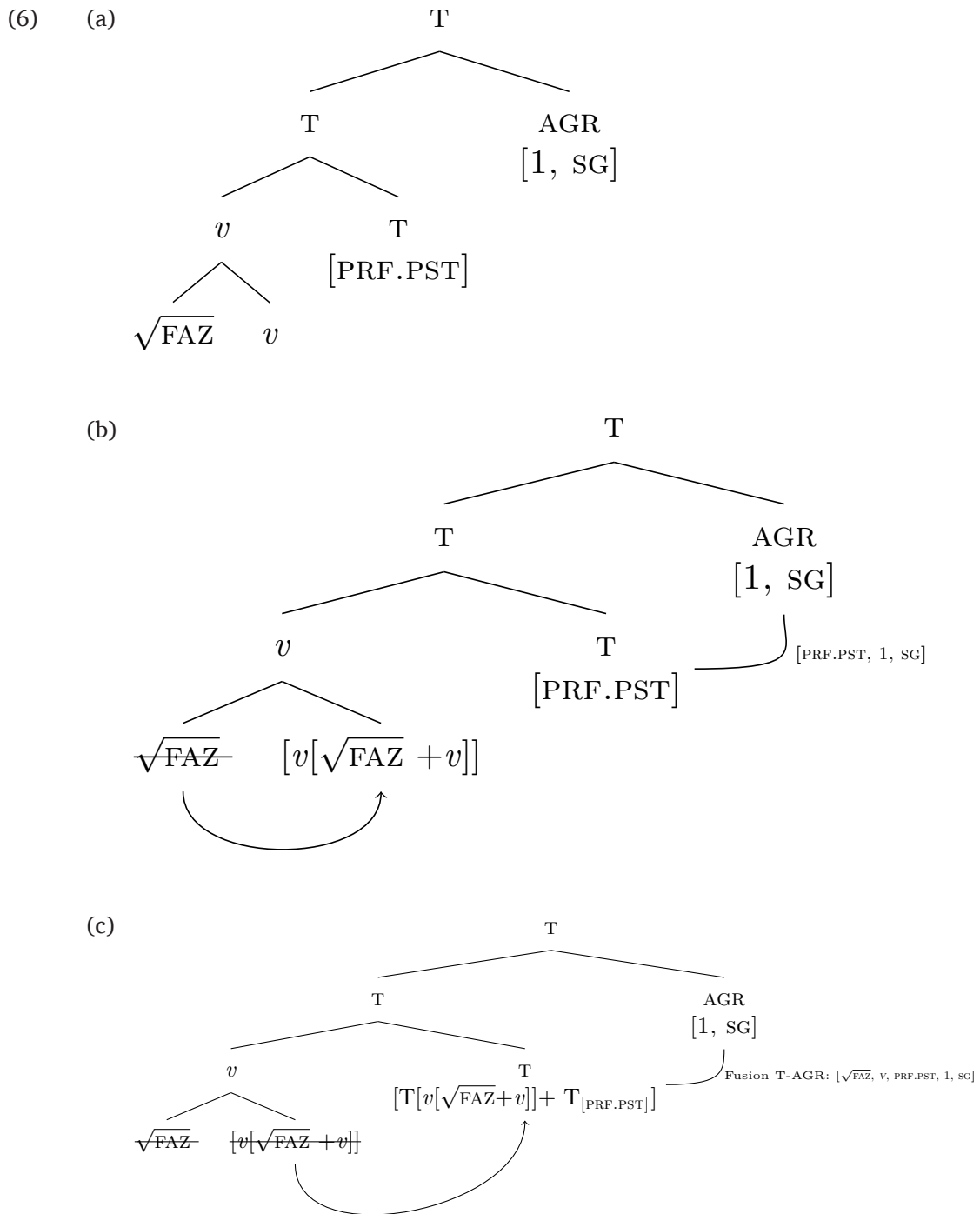
Table 2: Constraint *tableau* to generate overregularised form *trazi* (Lorandi, 2007, p. 156).

| Base | Candidate | OOROOTFAITH | ~OOROOTFAITH |
|--------|-----------|-------------|--------------|
| a) faz | ☞ fazi | | * |
| b) faz | fiz | *! | |

Table 3: Constraint *tableau* to generate overregularised form *fazi* (Lorandi, 2007, p. 158).

The *tableaux* above would describe the analysis made by the child: taking $\sqrt{\text{FAZ}}$ or $\sqrt{\text{TRAZ}}$ as a base, the best candidate is the one whose root fidelity is preserved, having as output *fazi/trazi*. According to the author, the adult grammar would be the one whose ~OOROOTFAITH restriction is higher in the hierarchy, which implies the ruling out of *fazi/trazi*, and accordingly, the production of *fiz/trouxe*.

Another proposal is Takahira’s (2013). The author assumes Distributed Morphology, which regards the atoms of the syntactic component as not being words but features. Thus, those features receive phonological content later (Late Insertion), following syntactic derivation. Also, some nodes, such as the Tense node and the Agr node, might fuse. Based on Siddiqi (2009), she proposes that when children overregularise, they do not move the verb to the T head as is necessary, as in (6a) and (6b), such that there is no complex head that comprises the root, the verbaliser, the tense morphemes when the Fusion between T and Agr occurs, as in (6c), the adult’s grammar.



In a nutshell, the changes from (6b) to (6c) require that one applies the Minimise Exponence (ME) constraint, which states that grammar realises the features of a given derivation in the most economical way, that is, the structure derived is the one that receives the least morphemes (Vocabulary Insertion) with the most features:

(7) MINIMISE EXPONENCE (Siddiqi, 2009):

The most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes.

Turning to the acquisition of irregular verbs, in Takahira's proposal, the child has not entirely mastered ME, letting Vocabulary Items be inserted separately: /faz/ for [$\sqrt{\text{FAZ}} + v$] and /i/ for the features [PRF.PST, 1, SG], after the Fusion of T and Agr. On the other hand, when adult grammar is achieved, and the movement from v to T takes place, Morphology will receive the instruction to spell out as /fiz/ the set [$\sqrt{\text{FAZ}} + v + \text{PRF.PST, 1, SG}$], respecting ME, since that realises the most features using less morphemes (in fact, a single one).

A third proposal (Wuerges, 2014) attempts to link verb input frequency and the probability of it being overregularised. Wuerges assumes Yang (2002), whose Variational Model of Language Learning offers a framework for understanding overregularisation as the misidentification of a verb as belonging to the wrong category. Taking a step back, according to the Variational Model, the probability of a certain rule being applied to a given verb depends on two other probabilities: that of the verb being identified as belonging to a given class and that of the rule itself being applied. In other words, verb membership to one of the irregular subclasses or to the regular class is regulated by how frequently this verb appears in the input, meaning that the child has more opportunities to assign the verb to its actual class. Of course, and this will be important down the road, in Yang's model, this is a probabilistic, rather than categorical, process. Different hypotheses of a verb's class membership may have similar weights in the early stages of acquisition. This predicts that verbs with different relative frequencies in the input, but belonging to the same rule class, may therefore be prone to disparate levels of overregularisation.

Wuerges (2014) explores this relative difference in her work by first identifying all irregular verb classes present in the language data directed to a single child. For example, the second-class irregular verbs *querer* 'want' and *dizer* 'say' belong to the same third-person singular present tense class: they are inflected with a zero morpheme, being realised respectively as *quer* 'he/she/it wants' and *diz* 'he/she/it says'. However, the former verb is much more frequent in the input than the latter (2029 *versus* 26 occurrences). This leads Wuerges to hypothesise, by looking only at the input frequency, that *dizer* will be much more prone to overregularisation, since the child takes longer to correctly assign it to the same rule class as the more frequent *querer*. This is in fact verified in the child's production. Hence, as the child learning their language comes into contact with more data, their hypotheses (implemented as probability weights) about a verb's class membership are increasingly refined, leading to less and eventually no overregularisation. At the same time, very frequent verbs should see less cases of overregularisation, as attested by Wuerges.

We have shown so far some facts about the acquisition of irregular verbs, in which children overgeneralise a rule for regular verbs. Additionally, we have seen three existing formal proposals to deal with those facts, Lorandi's (2007, 2010), Takahira's (2013) and Wuerges' (2014). The first, indeed, explains why children produce overregularised irregular verbs: it is a way of maintaining the verbal root, so the child, at first, would rather preserve the root than demolish it. Later, they would notice that demolishing it is the actual Portuguese grammar. However, this proposal does not account for the U-shaped curve, that is, why children first hit (stage 1), then miss (stage 2) and finally hit (stage 3) the verbal target again. Additionally, the root fidelity restriction ends up being a *tout court* principle, in which either the child has it or does not have it. On the contrary, the child does not categorically make mistakes during the U-shaped period, as can be seen in **Figure 1**, which would not be predicted if the root fidelity restriction was on the right track to account for those facts.

On the other hand, Takahira's (2013) proposal predicts that children during the period of overregularisation do not follow ME and, accordingly, do not present verb movement to T, which triggers vocabulary insertion to realise the root + *v* as /faz/ and the Tense/Agree as /i/, hence /fazi/. Again, this proposal fails to acknowledge the entire acquisition process of irregular verbs and, notably, expects that the child does not follow ME at all. In this case, there are questions in need of an answer: since the ME is not acquired, why do children make hits in the first place? Why does the child first derive a structure that requires ME? Is this ME absent in the entire child's grammar? We know, as just mentioned, that during the U-shaped process, children do not make mistakes categorically; they do make some hits; thus, that would imply that ME is sometimes in place and sometimes is not, which seems *ad hoc*. Finally, the main factor that triggers the non-realisation of most minor features is the absent movement of *v* to T. If one desires to corroborate Takahira's proposal, the task is to find evidence of the non-movement of the verb in Brazilian children's grammar. It follows that studies on verb movement in early childhood can shed light on the plausibility of the proposal, as mentioned earlier.

For instance, Santos and Lopes (2017), discussing the first steps of syntax acquisition, show that children do move *v* to inflection by an early age, since *v* is in a particular position to some adverbs, and VP ellipsis is found at the early stages of the acquisition process. VP ellipsis has been systematically related to verb movement to a position above VP,² where all the material that was below the verb in its first-merged position is elided, whilst only the V in a higher position is pronounced (e.g. question: “*Você deu o presente para a Maria?*” ‘Did you give the gift to Maria?’ / Answer: “*Dei.*” ‘I gave’, where [IP *dei* [VP *dei o presente para a Maria*]]). Armed with that assumption, the authors claim that the following VP ellipsis data indicates the earliest

² For a discussion on VP ellipsis and how this phenomenon is related to verb movement, cf. Cyrino and Matos (2002, 2011), Cyrino and Lopes (2016), Santos and Lopes (2017).

evidence that one can find to attest to verb movement to the inflection zone not only in Brazilian (8a) but also in European Portuguese (8b).

(8) (a) *Tomou remédio também?* (Adult)

took medicine too
 ‘Have you taken the medicine too?’
Tomou. (Child – 2;1)³
 took
 ‘I’ve taken it.’

(b) *O cavalo vai papar?* (Adult)

the horse will eat
 ‘Will the horse eat?’
Vai. (Child – 1;9.14)
 will
 ‘It will.’

With regards to the placement of some adverbs in relation to the verb, assuming some aspectual adverbs are in the inflectional zone and must be adjacent to the verb, Lopes (2009, p. 124) shows that the child, from an early age, does move the verb to the inflectional area, as in (9a) and (9b). Additionally, if one adopts Cinque’s (1999) proposal that all adverbs are in a fixed position in the inflectional area, the example in (9c) is further support to verb movement in a very early stage of the child’s grammar, since the verb is to the right of the fixed-order adverb *ainda* (‘still’). Note that if the verb did not raise, sentences such as **Não ainda cameu*, ungrammatical in the target grammar, would be predicted, in contrast with (9c).

³ It is worth noting that the evidence presented by Santos and Lopes (2017) is not from the ‘one-word’ stage. For instance, the Brazilian child from example (8a) produced more than one word at 2;1 and even earlier, at 1;10 (data from CEAAL/PUCRS – *Centro de Aquisição e Aprendizagem da Linguagem*).

(i) *eu tenho o pe(r)fume.* (2;1)

I have the perfume
 ‘I have the perfume.’

(ii) *é da bonequinha.* (1;10)

is from.the little.dool
 ‘It belongs to the little dool.’

Additionally, the child from example (8b) also produced more than one word before the age indicated, as Santos’ (2009, p. 215) example indicates.

(iii) MAE: *onde está o barco filho?*

where is the boat son
 ‘Where is the boat, son?’

*TOM: *ã [ː nãõ] s(e)i.* (01;06.18)

neg know
 ‘I don’t know.’

- (9) (a) *Aqui já comeu.* (= *Aqui (o boneco) já comeu*) (2;3)
 here already eat.PST (= here the doll already ate)
 ‘Look, it has already eaten.’
- (b) *Já tem out(r)o bicho.* (2;3)
 Already have.PRS another bug
 ‘There is already another bug.’
- (c) *Não comeu ainda.* ([IP *comeu ainda* [VP *comeu*]]) (2;3)
 no eat.PST yet
 ‘(He) has not eaten yet.’

The data just offered advocates against Takahira’s (2013) proposal in two ways. First, accepting her proposal requires assuming children do not move their verbs at all, contrary to the licensing of VP ellipsis, which demands verb movement and the placement of the verb in relation to aspectual adverbs from an early age: the period attested where children elide their VP and place their verbs to the left of adverbs (1;9/2;3) overlaps with the period where they overregularise. Second, Takahira’s (2013) proposal would only explain the second stage of the overregularisation process: BP children would move their verbs to the inflection zone at first, generating the target form (the first stage), and, oddly, they would disorganise their grammar, as though they had forgotten about how to move verbs, accounting for the second stage; finally, they would, again, remember and tune their grammar to properly move verbs, yielding the target form once again (the third stage). Under both points of view, Takahira’s (2013) proposal fails either based on the empirical evidence or on the logical claim requiring an *ad-hoc* account.

Wuerges (2014) will provide us with some important tools regarding the non-categorical nature of overregularisation, while being flawed in a way recognised by Yang (2016) himself. Let us first look at her work’s virtues. By assuming the Variational Model (Yang, 2002), this proposal is well-suited to make verifiable claims about children’s productions based on the language directed to them, while also conceding that different rules and verb class memberships may have competing probabilities at the earlier stages of acquisition. Consider again the child learning that *dizer* ‘say’ is an irregular verb. Since this verb is less frequent than others belonging to the same irregular third-person-singular-present-tense class (in the case of the child studied by Wuerges) there will initially be some non-zero probability that the child might classify *dizer* as a regular verb, applying its associated rule and producing /dizi/ ‘he/she/it says’. The very next time the child produces the third person singular present tense of *dizer*, however, there is also a chance that they will correctly retrieve its irregular membership, thusly producing /diz/. On to Wuerges’s shortcomings, which are really Yang’s (2002). In his 2016 book, Yang recognises that his previous model is “well equipped to handle noise – and only noise. It does not have the appropriate mechanism for distinguishing noise from exceptions.” (p. 6) To take from Yang’s example, if one percent of the time children hear linguistic data containing verb raising

in questions (“Baa baa black sheep have_t you ____t any wool?”), they will produce sentences with verb raising with a very low, one-percent chance. This is not, however, the correct English grammar: verb raising is an exception – a non-productive corner of the grammar – restricted to very specific contexts. Similarly, adults are known to make performance mistakes – a slip of the tongue – that look very much like overregularisations. Children must be equipped to overlook such noisy data in order to appropriately assign irregular verbs to their exception(al) classes. Otherwise, there would be a very small, although real, chance that overregularisations become a plausible, if only very rare, part of their grammar. The Tolerance Principle (Yang, 2016) represents an enhancement to this defect in the 2002 Variational Model. We adopt this more recent version of the theory, leading us to partly reject Wuerges’ work.

In the following section, we present the theoretical framework upon which our proposal is built. These are Distributed Morphology, which provides an adequate description of verbal derivation, and two general learning strategies related to the third factor, namely Minimise Maximal Means (Biberauer, 2017) and the Tolerance Principle (Yang, 2016), which children might follow when acquiring the pieces that constitute the verbs of their language.

3. Theoretical Tools

3.1 Distributed Morphology

The Minimalist Program (MP), motivated by trying to minimise the principles and operations that were only justified internally to Generative-Transformational Grammar but relatively unnecessary for the healthy functioning of the computational system, operates with a single basic syntactic mechanism – Merge – genetically circumscribed to the human species. For a large portion of the theories associated with the program, known as lexicalist, such operations are activated by features present in Lexical Items (LI). This means that words, even though they are not stored in the Lexicon entirely derived/inflected, are formed in a separate, lexical generative component. On the one hand, the weak Lexicalist hypothesis (Di Sciullo & Williams, 1987) is known for postulating that only derivation is the competence of the Lexicon, whilst inflection, of syntax. On the other hand, the strong lexicalist hypothesis (Anderson, 1982) assumes that both derivation and inflection are generated in the lexicon. Regardless of the hypothesis, the fact is that, in Lexicalism, two generative components end up being postulated: the Lexicon and syntax itself.

This paper adopts an alternative hypothesis to lexicalism that does not postulate any morphological module before syntactic derivation, a model known as Distributed Morphology (Halle & Marantz, 1993), or DM. By assuming that there is syntax throughout the derivation, reaching the interior of what we commonly call a word, DM considers that its primitives are subject to the exact mechanisms, restrictions and operations of the Minimalist Program: to cite a few, it assumes a binary branching structure with local and cyclic constraints, and with the basic operation, namely, Merge.

In this approach, what comes to be a word is the by-product of different syntactic, semantic and phonological features distributed in three lists, namely, List 1, or the Strict Lexicon, containing features to be manipulated by syntax; List 2, or Vocabulary, which contains pairings of phonological features to morphosyntactic features (Vocabulary Insertion) and, finally, List 3, or Encyclopaedia, which interprets the structure from a specific contextual instruction.

Thus, the syntactic derivation operates with items without any phonology (cf. Halle & Marantz, 1994), whose “sound” material is inserted post-syntactically, following the pairings found in List 2. This property of late insertion, along with the underspecification of features and syntactic structure throughout the derivation, differentiates DM from other (word-based) morphological theories within the framework of the Generative Grammar.

To illustrate how DM works, based on Bassani and Lunguinho’s proposal (2011), we will derive the verbal form *cantávamos* (‘sing.IPFV.PST’). A derivation under DM starts with selecting items from the Strict Lexicon, an inventory of (bundle of) features, roots, and categorisers. This selection feeds into Numeration (N), as in (10a), the essential ingredients to run a given syntactic derivation. For instance, for *cantávamos* (‘sing.IPFV.PST’), the items compounding N are: a ν verbaliser, a T abstract temporal morpheme specified as past imperfect⁴ and $\sqrt{\text{CANT}}_{[\text{cl1}]}$. The first step of the derivation is to select $\sqrt{\text{CANT}}$ and the ν categoriser, merge them, and decrease their indices to 0 in N. Next, the T bearing [IPFV.PST] feature is inserted, being merged with object X. That yields the structure in P, represented in (10d).

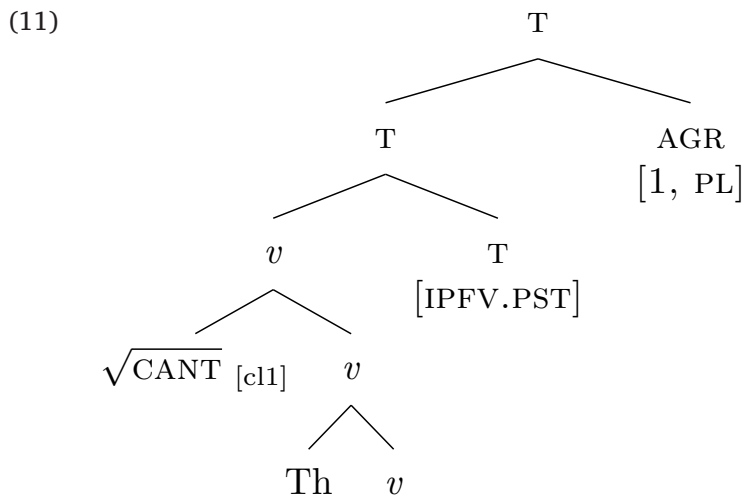
- (10) (a) $N_0 = \{\nu_1, \sqrt{\text{CANT}}_{[\text{cl}_1]}_1, [\text{T}_{\text{IPFV.PST}}]_1\}$
 (b) $N_1 = \{\nu_0, \sqrt{\text{CANT}}_{[\text{cl}_1]}_0, [\text{T}_{\text{IPFV.PST}}]_1\}$
 Select and Merge = $[\sqrt{\text{CANT}}_{[\text{cl}_1]}]$ and $[\nu]$
 $X = [[\sqrt{\text{CANT}}_{[\text{cl}_1]}] [\nu]]$
 (c) $N_2 = \{\nu_0, \sqrt{\text{CANT}}_{[\text{cl}_1]}_0, [\text{T}_{\text{IPFV.PST}}]_0\}$
 Select $[\text{T}_{\text{IPFV.PST}}]$ and Merge with X
 $P = [[[\sqrt{\text{CANT}}_{[\text{cl}_1]}] [\nu]] [\text{T}_{\text{IPFV.PST}}]]$
 (d)
-
- ```

graph TD
 T1[T] --- nu1[ν]
 T1 --- T2[T]
 nu1 --- sqrtCANT["√CANT [cl1]"]
 nu1 --- nu2[ν]
 T2 --- IPFV["[IPFV.PST]"]

```

<sup>4</sup> We warn the reader that Tense and Aspect may be separate heads in the syntactic derivation, such as proposed by Ippolito (1999). However, for the purpose of our discussion, this split is not relevant.

When the indices in N have been exhausted, the structure is Spelled-Out.<sup>5</sup> It is interpreted in Logical Form and, after going to the Morphological Structure (MS), it receives phonological content. In MS, as a morphological well-formedness condition, an AGR node is inserted in T, at the same time as a thematic node, hosting a verbal thematic vowel, is inserted in the *v* categoriser (cf. Harris, 1999; Ippolito, 1999; Oltra-Massuet, 1999). The structure, then, becomes (11) after the insertion of the dissociated nodes – that is, nodes without syntactic-semantic relevance.



The terminal nodes from (11) receive phonological content from List 2 at Vocabulary Insertion. Within this list (cf. (12)), for example, it is specified that for the  $[cl_1]$  class feature born by the root, the /a/ theme vowel will be inserted. In the case of *beber* ‘drink’, bearing the  $[cl_2]$  feature, this is an /e/ theme vowel. Forasmuch as temporal imperfect past morphemes are different depending on the class a root bears (/va/ for class 1, as in *cant-a-va*; /ia/ for classes 2 and 3, as in *beb-ia*, *sorr-ia*), it is specified in the Vocabulary insertion that the /va/ morpheme is inserted in the context of [IPFV.PST] along with (represented by the bar ‘/’) the  $[cl_1]$  feature, whilst the [IPFV.PST] feature in the context of  $[cl_{2/3}]$  is realised as /ia/. Finally, /mos/ is inserted for the [1, pl] features born by the AGR terminal node.

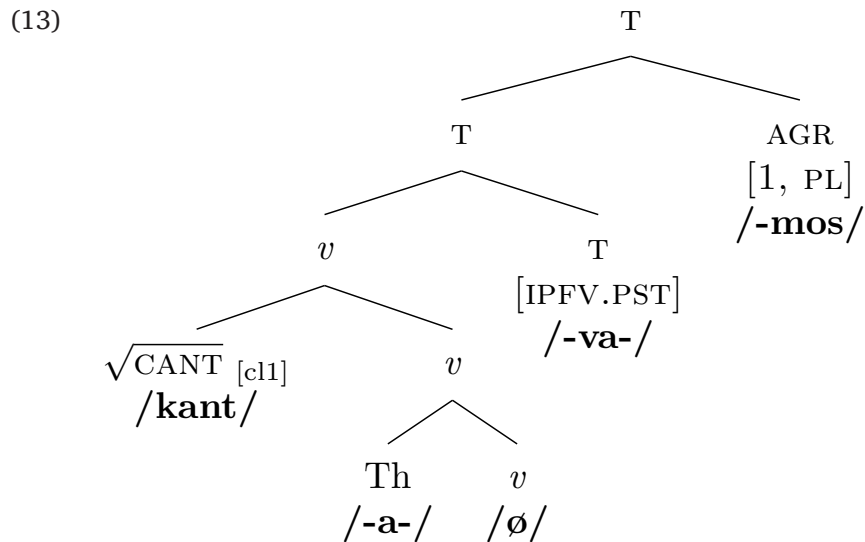
(12) Vocabulary Items

|                              |                                          |
|------------------------------|------------------------------------------|
| /a/ ↔ $[cl_1]$               | /ia/ ↔ [IPFV.PST] / $[cl_2]$ or $[cl_3]$ |
| /e/ ↔ $[cl_2]$               | /u/ ↔ [PRF.PST, SG]                      |
| /i/ ↔ $[cl_3]$               | /i/ ↔ [1, SG]                            |
| /va/ ↔ [IPFV.PST] / $[cl_1]$ | /mos/ ↔ [1, PL]                          |

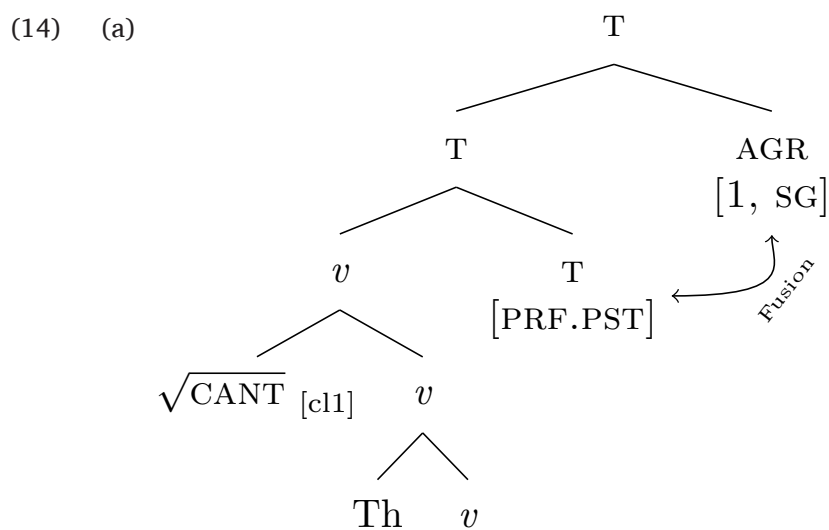
<sup>5</sup> Although *v* is taken to be a phase head (Marantz, 2001), we understand along with Chomsky (2001) that a phasal domain is Spelled-Out only after the merge of the following phase head, not upon the merge of the phase head itself. In the case of (10d), the next derivation step that sends *v* phase domain to spell-out is upon the merge of C, a category that we do not represent here (cf. also Embick 2010).

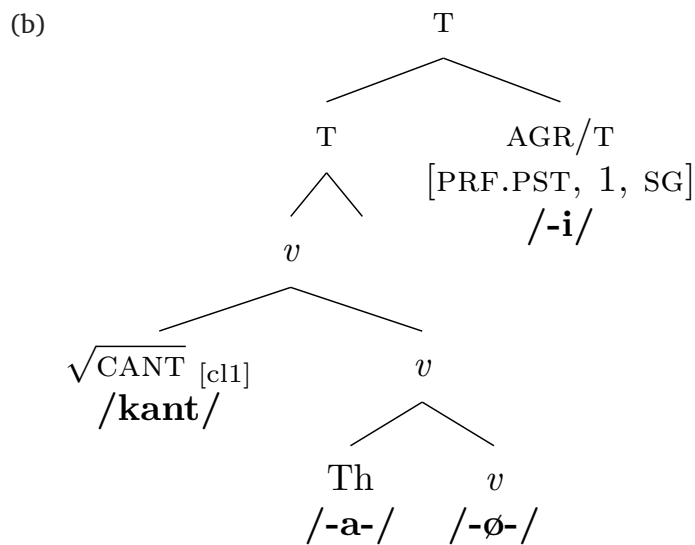


With the vocabulary items paired with their respective bundle of features, the structure is operable by the phonological component, giving it the appearance (13) of a word such as we know it.



Having illustrated how DM works, we can now turn to the derivation of the Portuguese simple past, from which we delineate our proposal. There are some cases in which, within the morphological structure, the post-syntactic Fusion operation occurs, joining functional nodes. Bassani and Lunguinho (2011) show this is the case for the perfect past bearing any combination of person and number features. For example, for the perfective past form *cantei* ‘I sang’, the structure that arrives in MS is (14a); the structure after the Fusion operation between T and AGR nodes and vocabulary insertion is illustrated in (14b).





Given the Vocabulary items in (12), the phonological pairing of the [PRF.PST, 1, SG] features could be either /i/ ↔ [1, SG] or /u/ ↔ [PRF.PST, SG]; however, the first, Bassani and Lunguinho (2011) propose, has insertion priority, following the fact that the first person is more marked, as suggested by Galves (2001). The derivation, however, is not complete with the Vocabulary Insertion since the verbal form generated at this moment is /cantai/ and not the goal form /cantei/. Here, a phonological rule is applied, due to the allomorphy in these segments, raising /a/ to /e/ in the context of /i/: hence, /cantai/ is realised as /cantei/ (cf. Bassani & Lunguinho, 2011).

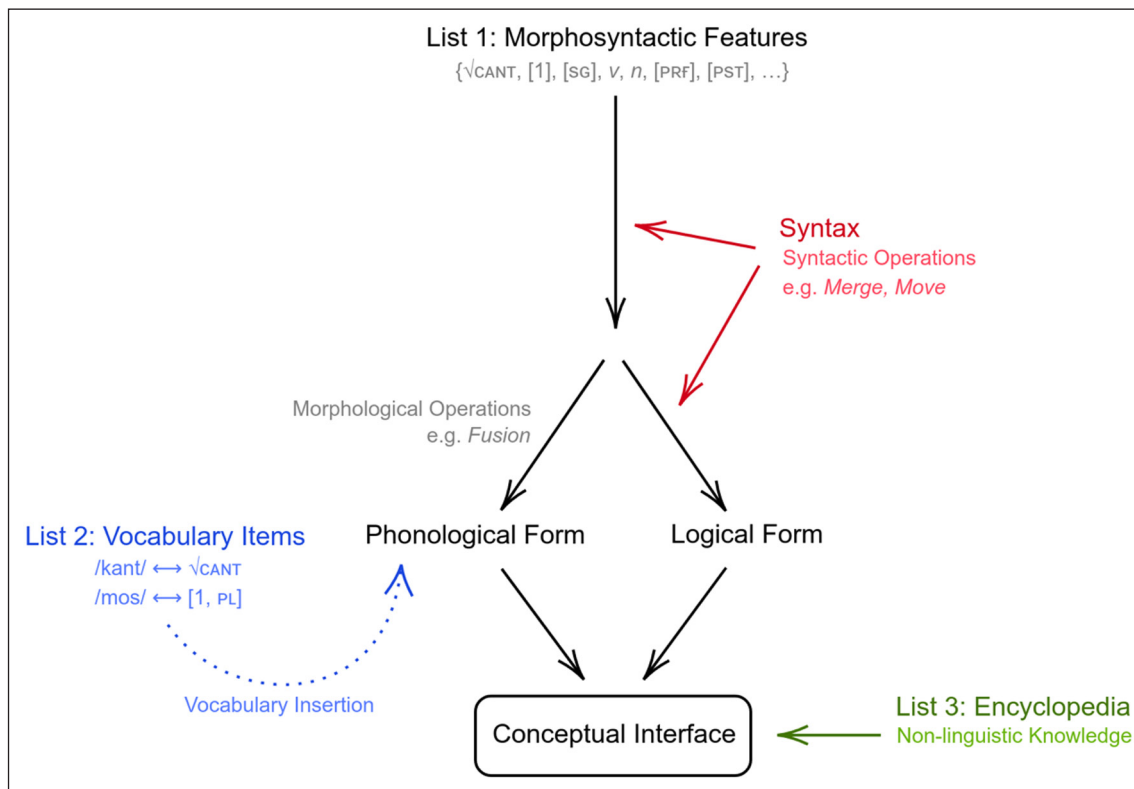
Note that under the DM approach, words such as *cantávamos* or even the apparently more atomic *flor* ‘flower’<sup>6</sup> are not an indivisible (syntactic) unit. They are, on the contrary, a complex structure, yielded from (i) syntactic operations, such as Merge, (ii) morphological operations, such as of Late Insertion, Fusion, and (iii) phonological operations, such as Vocabulary Insertion and allomorphy rules. For instance, *flor* is minimally composed of the root √FLOR and a noun categoriser *n*. Under this background, what is known as a word is spread (or “distributed”) throughout various lists within the architecture of grammar (cf. **Figure 2**), and not found in a single place, namely the Lexicon, as a proponent of lexicalism would argue.

As for language acquisition in Generative Grammar, since its earlier stages (Chomsky, 1965, 1988), this theory tries to understand how children learn a language rapidly, even though they face significantly incomplete and degraded linguistic data. This puzzle, known as Plato’s Problem, lead Chomsky to hypothesise a biologically innate language acquisition device (LAD), called Faculty of Language (FL). From FL, general language principles, previously set, along with the

---

<sup>6</sup> We thank the anonymous reviewer for the suggestion of the word *flor*, which we have decided to adopt.

setting of language-specific properties, the parameters, compound a Universal Grammar (UG), the initial stage of FL.



**Figure 2:** The architecture of Grammar according to DM. Based on Sidiqqi (2009, p. 14).

The Principles and Parameters Theory was the solution proposed for the fact that the language acquisition process occurs in a small window of time, creating order out of linguistic chaos – Plato’s Problem. Thus, acquiring a language was seen as nothing more than a natural process where UG is fed by linguistic stimuli, the input, to set properties specific to the target language – the parameters. In this vein, the setting of parameters would be a mechanism that simplifies the task of acquiring languages: language acquisition is simply explained through the richness of UG in terms of the content that is useful to lead to adult grammar. Nonetheless, with the emergence of the Minimalist Program (Chomsky, 1995), FL starts to be considered a perceptual and broader cognitive system, not assigning the totality of linguistic properties to UG. In this sense, for Chomsky (2005), acquiring a language is not dependent only on UG, but on the interaction of three factors, namely (i) UG itself, (ii) primary linguistic data (PLD) and (iii) computational efficiency principles, such as Feature Economy, Input Generalisation (cf. Biberauer, 2019a,b) and the reinforcing/punishment of a structure that is (not) equivalent to the

adult grammar, as is proposed by Yang's (2016) work on how exceptions play a role in informing the child of whether a rule is productive or not.

Within this not-that-rich approach to UG, the answer to Plato's Problem is one more time invoked, and the third factor of this new language design (Chomsky, 2005) seems to solve the question that backtracks to the earlier stages of Generative Grammar: "How should a theory of grammar simplify the learner's task to achieve successful acquisition with a relatively small quantity of data?" (Yang & Roeper, 2011, p. 560). Having outlined the theoretical background we are assuming for grammar, in the next section, we present the three factors in language design, focusing on the third one, which we will illustrate with Biberauer's (2011) maximisation of features, and Yang's (2016) mathematical model of language acquisition.

### 3.2 Some general learning computational strategies

In the pursuit of a leaner model of the core aspects of grammar, Chomsky (2005, p. 6) claims that what is behind the "growth of language in the individual" is that "the faculty of language has the general properties of other biological systems", known as the *three factors*,

- (i) **Genetic endowment** (apparently nearly uniform for the species, which interprets part of the environment as linguistic experience, a nontrivial task that the infant carries out reflexively, and which determines the general course of the development of the language faculty.)
- (ii) **Experience** (which leads to variation, within a fairly narrow range)
- (iii) **Principles not specific to the faculty of language**

The first of which concerns the innate UG, which would comprise mechanical operations, such as Merge, Move (still [internal] Merge), and, at least at first sight, Fusion (within DM). We take Fusion and any other essentially morphologic operation to be the expression of the first factor given that morphology is arguably exclusive to language. The second is related to the Primary Linguistic Data that function as input/intake. Finally, the third factor is closely related to efficient computational strategies during the acquisition process, which, according to Chomsky (2005, p. 6), include subtypes of third-factor principles:

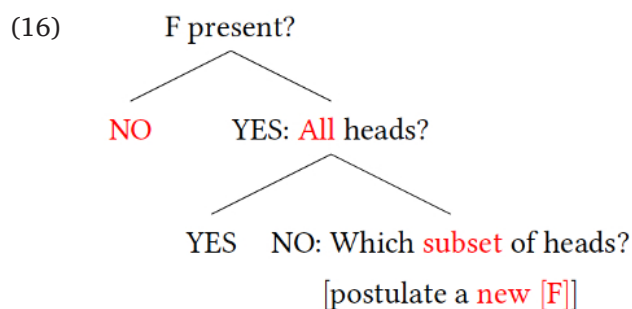
- (a) principles of data analysis that might be used in language acquisition and other domains;
- (b) principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including principles of efficient computation, which should be expected to be of particular significance for computational systems such as language. It is the second of these subclasses that should be of particular significance in determining the nature of attainable languages.

Thus, as with any biological system, the faculty of language is aided by factors other than its core functioning and the data available, that is, third-factor aspects that end up easing the task

of acquiring a language: it is much less costly for a child to postulate as few formal features (FF) as possible to account for the linguistic stimuli and in doing so, further economic to maximise those already hypothesised features. This new approach for language acquisition where the child follows the most economical strategies promotes general learning biases as a guide to achieving the target grammar. Such a proposal is Biberauer's (2011) MAXIMISE MINIMAL MEANS, which integrates Feature Economy (Roberts & Roussou, 2003) and Input Generalisation (Roberts, 2007).

- (15) (a) Feature Economy (FE): Postulate as few FFs as possible, given the PLD.  
 (b) Input Generalization (IG): Maximize available FFs.

The conjunction of FE and IG yields a path for language acquisition. The acquirer first postulates the NO feature as a default setting (Biberauer, 2017; Roberts, 2019), respecting both FE and IG. Later, in having identified a feature, the child generalises it for ALL heads of the relevant type – once postulating a feature, they violate FE, since hypothesising one is worse than none, but obey IG, given the extension to all categories related. Finally, if a set of heads is not identified as bearing such a feature, the procedure stops and understands that this feature is not present in all heads but only in SOME. This NO > ALL > SOME learning path can be depicted as follows:



Within this discussion, exposure to PLD, specifically the intake (Evers & Van Kampen, 2008), is essential to steer the learning path presented above. Then, being exposed to enough sentences with the relevant and prominent features, the child can start their journey. Those sets of “sentences” are defined as *triggers*, which consecutively are correlated with *formal feature expression* ( $FF_i$  – a piece of input text in which a FF is present), as represented below.

- (17) Trigger: a substring of the input text of the PLD. S is a trigger for an optional  $FF_i$  if S expresses  $FF_i$ .

In sum, those definitions state that a trigger that leads children to set a particular language-specific property is the linguistic stimuli that include the relevant evidence from which they set their target grammar. Thus, in other words, as Roberts (2012, p. 321) claims, acquiring a language, under this approach, is the searching for the easiest setting compatible with the PLD:

The acquisition device searches the space by looking for the ‘easiest’ solution at each stage, where a solution is defined as a parameter-setting compatible with available primary linguistic data. The device moves from a relatively easy to the next-hardest stage only when forced to by primary linguistic data (PLD) incompatible with the current setting.

Still, an important question regarding language acquisition is: *how* does a child move from one solution to another? We want to suggest that the steering of language acquisition as the moving down the (learning path) tree (cf. (16)), in Roberts’ (2012) words, is accounted for by Yang’s (2016) Tolerance Principle, another principle of computational efficiency.

Languages can be designed in one of two ways: either there is a core grammar, responsible for generating licit phrases and sentences, and a periphery (Chomsky, 1981, p. 8) where the unruly aspects of that language are stored; or, as some suggest (e.g., Tomasello, 2005), all possible constructions, ranging from the very concrete to the very abstract, are stored. Since Generative Grammar assumes the first position, its proponents are left with the hairy problem of explaining how a child, having set a parameter, does not overgeneralise it to the grey zones of language where it should not apply. For example, in Portuguese, adjectives generally come after a noun head: *uma casa amarela* ‘a yellow house’, *um dia longo* ‘a long day’. However, a few adjectives can precede the head of the noun phrase, as in *um grande homem* ‘a great man’. If such forms are postulated to be stored, then the whole theory becomes vulnerable to the slippery-slope argument that *all* linguistic forms could be stored as well. A clear line must be drawn to separate the parts of the language which are generated and those which are committed to memory.

Inductive methods of learning grammar resistant to the noisy aspects of linguistic input have already been proposed under the idea that different grammars compete until a stable one is achieved (Yang, 2002). Nonetheless, such methods are still subject to what Yang (2016) calls the “leaky” parts of grammar. For example (pp. 42–43), a rule-finding algorithm<sup>7</sup> could learn the *-ed* rule of past formation in English by generalising over successive pairs of linguistic data. Let us start with the pair *walk-walked*, which leads the learner to entertain the rule IF walk THEN -d, meaning, “if you want to produce the past of *walk*, then add the *-d* to its end”. The next piece of data, the pair *talk-talked*, prompts the algorithm to generalise the previous rule further, yielding the rule in (18):

(18) IF \*alk THEN -d

This can ultimately lead the algorithm to discover the (correct, productive) past tense rule

---

<sup>7</sup> An anonymous reviewer observed that omitting part of the root might be incompatible with a Distributed Morphology approach. Although this is true, this rule-finding algorithm is described as a way of illustrating a plausible learning mechanism that might be employed by children. Yang (2016) does not commit to it and neither do we. Further research is needed to determine which hypothesis-generating mechanisms would be compatible with our current proposal.

(19) IF \* THEN -d

which is productive for all but, ironically, the most frequent English language verbs, which is top-heavy in terms of irregulars. Of course, the fact that children overregularise (Marcus et al., 1992), as discussed in the beginning of this paper, constitutes evidence in favour of the rule-finding algorithm just described (presented by Yang, 2016, based on Yip & Sussman, 1997); but it still does not account for how actual children eventually learn the correct form. Incidentally, the fact that this algorithm is always looking for further generalisations is an illustration of what Yang (2016, p. 72) calls Maximise Productivity (20), a guiding principle in language acquisition:

(20) MAXIMISE PRODUCTIVITY

Pursue rules that maximize productivity.

This principle is motivated by the fact that subsequent analyses of the data are necessary for the child to come up with a maximally productive rule. This is a way of avoiding having different rules for different verb forms, such as IF \*alk THEN -d in order to produce *talked* and *walked*, and IF \*rk THEN -d to produce *barked* and *worked* etc., when a more general rule would apply. In fact, Maximise Productivity eventually finds the most economical way of *processing and retrieving* linguistic information, within a psycholinguistic perspective. Note that an added bonus of Yang's model is that it captures *discrete infinity* (Chomsky, 1957, 1969). This way, new linguistic structures (children's overregularisations or morphological processes of word creation) are always accounted for.

Before moving on to the Yang's solution to the current conundrum of how children learn exceptions, it is essential to understand that the aforementioned algorithm could also lead to the curious situation in which *unproductive* rules are considered *productive*. The pairs *sing-sang* and *ring-rang* could be taken as evidence for another rule, namely (21) below. According to it, a verb such as *sting* should have its past form as *stang*, which is not the case. Interestingly, children do produce such overirregularisations – Marcus et al. (1992) attest cases such as *brang* and *wope* – although Yang (2016) argues that they are at least an order of magnitude less frequent than irregular verbs being regularised (0.02% overirregularisation against 4%–10% overregularisation, depending on the study).

(21)  $i \rightarrow \text{æ} / \_n$

One reason why overirregularisation is not observed very often may boil down to the fact that learning must occur within the reasonable processing time; that is, it is constrained by what is psycholinguistically efficient: “rules and exceptions are organised to optimise/minimise the time complexity of language use” (Yang, 2016, p. 60). Studies on comprehension time have shown (pp. 56–60) that exceptions tend to be processed faster than regular rules. One example is the processing of idiomatic expressions over expressions that have a compositional meaning, such as

*kick the bucket* (= die) vs. *lift the bucket*. In one study, the use of an idiom sped the reading up by about 100 ms (Swinney & Cutler, 1979). Yang then argues that such results are evidence that the Paninian Elsewhere Principle is a valid psycholinguistic model of processing and shines a light on how rules are processed.

The Elsewhere Principle, or Subset Principle (Berwick, 1985), states that more specific rules always precede the application of more general ones. Thus, the particulars of the psycholinguistic optimisation mentioned above are as follows: if a rule is productive, then its exceptions (or more specific rules) are listed and accessed first, in the fashion illustrated in its most basic form in (22) below.

(22) Rule  $R$   
 IF  $e_1$  THEN ...  
 IF  $e_2$  THEN ...  
 ...  
 IF  $e_n$  THEN ...  
 Rule  $R$ : IF \* THEN ...

For a rule  $R$ , its exceptions ( $e_1, e_2, \dots, e_n$ ) are listed according to their relative frequency, such that more frequent words are retrieved more quickly. If an item  $i$  does not fit any of these exceptions, then the regular rule is applied as a catchall condition. For irregular verbs, for example, the claim is that they are listed (and retrieved) somewhere before the regular rule. This explains why irregulars are never regularised in adult grammar, barring any processing time errors.

Naturally, the number of exceptions cannot be infinite, else the point of regular rules would be ruined. Thus, Yang (2016, p. 48) proposes a principle which “minimizes the computation of rules and exceptions”, dubbed the Tolerance Principle (23), which captures the precise balance between the psycholinguistically plausible timing and the implausibility for traversing the list of exceptions. In other words, it is a proportion limiting the number of exceptions a (putative) rule may endure before it is cast away as being unproductive.

(23) TOLERANCE PRINCIPLE

$$e \leq \theta_N \text{ where } \theta_N := \frac{N}{\ln N}$$

In the equation,  $N$  is the number of items fitting the structural description of a would-be rule.  $e$  is the number of exceptions to that rule. Thus,  $\theta_N$  is the threshold, or the number of exceptions that a rule can withstand and is defined as the product of the number of items divided by its natural logarithm. For instance, if there are ten candidates for a given rule, the number of exceptions tolerable is  $\theta_{10} = 4$ . One important property of this equation is that the higher the



number of candidates  $N$ , the lower the number of exceptions  $e$  that are tolerable before the rule crumbles.

Let us consider, as an example, a toy language consisting of six verb pairs (24). Further, assume our learner has encountered the data in the order they are presented.

- (24) (a) ming-mang  
 (b) bling-blang  
 (c) zing-zingged  
 (d) shing-shingged  
 (e) scring-scringged  
 (f) pling-plingged

Given the two first pieces of data, (24a,b), our learner might entertain a rule such as (25), reproduced below.

- (25)  $i \rightarrow \text{æ} / \_ŋ$

When the third piece of data (24c) enters their vocabulary, our learner will still consider rule (25) to be productive, listing *zing* as an exception (or potentially producing *zang*), since  $\theta_3 = 3$ . When *shing-shingged* (24d) is acquired, their vocabulary now has four items to which (25) could apply, two of which are exceptions. However, since  $\theta_4 = 3$ , the learner has no reason to throw the rule away. It is only when *scring-scringged* (24e) and *pling-plingged* (24f) join their vocabulary that the threshold  $\theta_5 = 3$  is exceeded since now there are four exceptions to rule (25). This triggers a re-evaluation of the grammar under consideration, leading to the deprecation of (25) and to the postulation of the new, correct rule (26):

- (26) Past Formation Rule  
 IF ming THEN mang  
 IF bling THEN blang  
 Rule: IF \* THEN -d

Since, in this language, the number of exceptions  $e$  will always be lower or equal to the threshold  $\theta_N$ , rule (26) above will never lose its productive status. Further, any new exceptions can readily be added to the list preceding the application of the rule.

The early period in which the unproductive rule (25) was considered to be productive appears to be analogous to what happens in English. Yang (2016, p. 84), based on child-directed speech data, shows that, since the most frequently occurring verbs are irregular, the earliest rule hypotheses available to the child are only those matching the structural description of said irregulars. This means there is a window of opportunity to observe overirregularisations, such as *blink-blank*. Still, due to its coinciding with the earliest stages of acquisition, memory and articulation constraints make observations of such productions rare to find, according to the

author. Furthermore, since different hypothetical rules to generate irregulars in English (e.g., *feed-fed*, *fly-flew*) are supported by the data, the period in which a particular unproductive rule is active turns out to be quite short. On the other hand, later in acquiring verbs, when the child has amassed around 1000 verbs total, the balance tips, and they enter the second phase of the U-curve, when overregularisations are indeed attested.

Having presented some general learning strategies that steer language acquisition, we can now explain the acquisition of the Brazilian Portuguese verb system by applying the ideas just exposed. In the next section, we piece together an analysis alternative to Lorandi's (2007, 2010), Takahira's (2013), and Wuerges' (2014) accounting for all three stages of the U-shaped curve.

#### 4. Alternative Analysis

As we have shown, the two formal analyses available to explain the U-shaped curve of irregular verb acquisition in Brazilian Portuguese, namely those of Lorandi (2007, 2010) and Takahira (2013), are incomplete descriptions of the development of child grammar. The main flaw in both proposals is that the principles/constraints that children seem to break in each case –  $\sim$ OOROOTFAITH and Minimize Exponence<sup>8</sup> (Siddiqi, 2009) – are presumably innate (since they are theoretical “principles”) and thus should be observed throughout the development of language in children. This does not mean that one should necessarily observe a principle's surface consequences from the outset of language production (cf. Bertolino & Grolla, 2012; Grolla, 2012, 2013); however, if that is the case, then the burden of proof (and indeed a convincing account) falls upon those who propose that a principle should not be observed at some given point in language acquisition.

The analysis by Wuerges (2014), on the other hand, proved promising in two ways (cf. section 2): it allowed the author to correctly predict which verbs would be more prone to overregularisation, based on input frequency, as well as being compatible with the fact children are non-categorical when overgeneralising. These merits are drawn from the Variational Model (Yang, 2002) which, as we have reviewed, is unable to deal with exceptions. We were led to partly reject her analysis on this basis. Further, we also diverge from her account in terms of the formal treatment she gives to irregular verb inflections. Following Yang, Wuerges proposes a set of phonological rules for inflecting regulars and irregulars. We adopt the view that all morphological processes are syntactic all the way down, as we explain below.

---

<sup>8</sup> Under Takahira's proposal (p. 436), Minimize Exponence is read as a principle of grammar and, thus, we understand it to belong to the first factor. At first sight, there is no reason not to consider Minimize Exponence as being a third-factor, principle like others cited in this paper, since it deals with how a derivation might run more economically (Siddiqi, 2009, p. 4).

Our alternative account of the U-curve follows Takahira's (2013) in her choice of Distributed Morphology as a descriptive background; however, we otherwise diverge from the very start. We claim that the three stages that can be inferred from the U-curve – high rate of target forms, low rate and high rate again – are explained solely by the differences in what morphemes are being listed by the child at any given time. In Stage 1, the grammar lists root/features-phonological content pairings for all verb forms, and thus we do not observe (nor expect) any overregularisations by the child. In Stage 2, enough data has been observed and features start being mapped to their phonological exponents, instead of being listed in each verb's entry. Finally, Stage 3 is the recovery of the balance within the child's Vocabulary.<sup>9</sup>

Let us explore this idea in some detail by characterising the child's knowledge at each stage. In Stage 1, we claim that in Brazilian Portuguese, all verb forms are memorised at first, until the inductive mechanisms described in our discussion of Yang (2016) above can generalise rules from accumulated linguistic input. To limit our discussion to past tense verbs – even though our approach should in principle extend to other irregular forms – children start by assigning sound to a complex comprising a root and features as in (27a–c). This association, which we have been calling a Vocabulary Item following DM, applies to regular and irregular verbs alike. Thus, the Vocabulary entries in a child's grammar in the early stages of acquisition (or Stage 1 of the U-curve) are a simple list, such as:

- (27) (a) /fiz/ ↔ [√FAZ, v, PRF.PST, 1, SG]  
 (b) /komi/ ↔ [√COM, v, PRF.PST, 1, SG]  
 (c) /dormi/ ↔ [√DORM, v, PRF.PST, 1, SG]

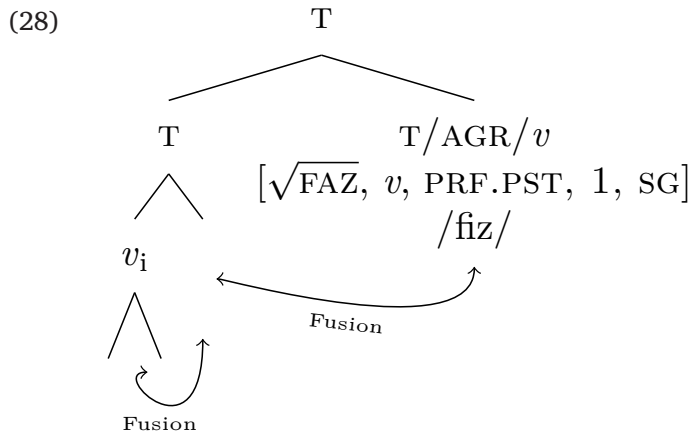
Listing (in terms of DM's Vocabulary) is the first step in acquiring the rules of past formation (Yang, 2016) and thus should not be controversial. Memorised chunks of language fitting some pattern are, in fact, the material for any rule-finding algorithm since they work by comparing data. Furthermore, the listing is necessary for dealing with exceptions once rules have been found.

An important consequence of early listing is that Siddiqi's (2009) Minimize Exponence (ME) principle does not need to be broken, as proposed by Takahira (2013), but rather triggers itself the Fusions represented by the arrows in (28). This leads to the late insertion of (27a), yielding the target irregular past perfect form *fiz* 'did'. An added benefit of this analysis is its compatibility with verb movement to T, which is attested in young children's grammar, as discussed in section

---

<sup>9</sup> An anonymous reviewer asked if there is evidence that children morphologically analyse verbal input and, conversely, how we could know if they are not just repeating forms they hear. We direct the reader to Resende (2021), who argues, based on empirical data, that very young children are sensible to the morphological makeup of words.

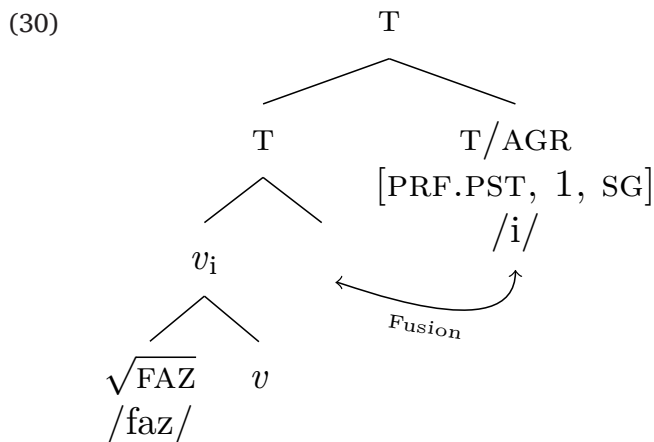
2. Apart from being useful in explaining the derivation below, since ME is (apparently a third-factor) principle, it is reasonable to assume that it should be followed from the outset of language acquisition, *contra* Takahira’s proposal.



We now move on to Stage 2, the dipping point of the U curve marked by the onset of overregularisations in a child’s productions. In terms of DM, a critical mass of Vocabulary Items has been collected, and the Vocabulary is reduced by generalisation algorithms (Yang, 2016). Vocabulary Items become more abstract if redundant features can be identified, extracted and assigned to a new morph. For example, suppose we accept (27) to be a reasonable approximation of Stage 1 knowledge of verbs. In that case, a generalisation algorithm could determine that the features [PRF.PST, 1, SG] should not be listed with the roots but rather as a morpheme of its own. This leads to the rise of (29d), the abstract past perfect first-person singular morpheme, and a list of different root (29a–c) morphemes.

- (29) (a) /faz/ ↔ [√FAZ]  
 (b) /com/ ↔ [√COM]  
 (c) /dorm/ ↔ [√DORM]  
 (d) /i/ ↔ [PRF.PST, 1, SG]

Notice that while the past tense verb *comi* ‘I ate’ and *dormi* ‘I slept’ can be generated by the Vocabulary above, the ability to generate *fiz* ‘I did’ has been lost. In its place, the form *fazi* ‘I doed’ is to be expected, as is illustrated in (30). This structure describes at the same time the derivation of regulars (*comi*, *dormi*, *corri* [‘I ate, slept, ran’] etc.) and irregulars (*boti*, *di*, *pensi*, *dobri*, *tomi* [‘I putted, gaved, thought, folded, taked’], Lorandi, 2010). Further, since the generalisation algorithm has deleted the exceptional (irregular) forms, Minimize Exponence is being dutifully respected.



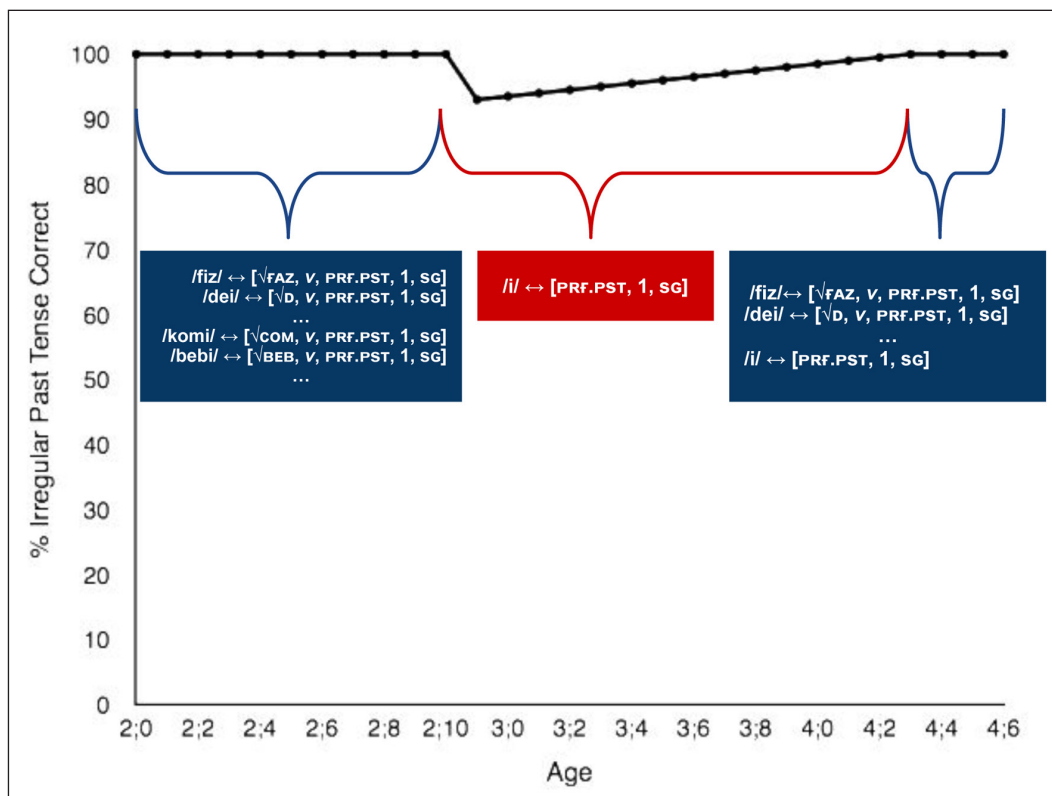
The gradual progression into Stage 3, in which target forms are regularly produced again, can be described as the settling in of exceptions. If the child has lost a VI such as  $[\sqrt{\text{FAZ}}, v, \text{PRF.PST}, 1, \text{SG}] \leftrightarrow /fiz/$  when entering Stage 2, a step into Stage 3 would be its resettling into their Vocabulary.<sup>10</sup> An important consequence of listing bundles of features, including a root, is that Minimise Exponence would trigger once again the movement of  $v$  into T, meaning that the target irregular form would be inserted as was illustrated in (28). This precedence of the most specific can be seen as an implementation of the Paninian Elsewhere Principle, which is paramount to Yang's (2016) argumentation for the listing of exceptions before regular forms. Within DM, the insertion of the best-fitting, most specified morpheme before more general ones captures this. Once all irregulars have been properly identified and listed (cf. (31)), the child leaves behind Stage 2 and reaches Stage 3, which is tantamount to adult knowledge of verb tenses. A summary of all three stages can be seen in **Figure 3**.

- (31) Past-formation rule (first person singular verbs)
- $/fiz/ \leftrightarrow [\sqrt{\text{FAZ}}, v, \text{PRF.PST}, 1, \text{SG}]$
  - $/dobr/ \leftrightarrow [\sqrt{\text{DOBR}}]$
  - $/dei/ \leftrightarrow [\sqrt{\text{D}}, v, \text{PRF.PST}, 1, \text{SG}]$
  - ...
  - $/i/ \leftrightarrow [\text{PRF.PST}, 1, \text{SG}]$

Two points should be addressed before concluding this section. Firstly, as mentioned before, children do not categorically produce overregularisations in Stage 2. Rather, such mistakes happen at a rate varying from 4% (Marcus et al., 1992) to 10% (Yang, 2016) of all verb productions. If a child's Vocabulary underwent a thorough revision upon entering Stage 2,

<sup>10</sup> As discussed elsewhere and mentioned later in this section, in Stage 2, the production of overregularisations is not categorical. This is where Wuerge's (2014) and consequently the Variational Model (Yang, 2002) come into play. These approaches are able to explain why only a fraction of irregular productions are deviant: there are competing rules, each with a non-zero probability of being activated, being considered in parallel by the child.

such that all redundancies were broken down to morphemes of their own right, then it would stand to reason that all verbs would be derived following the derivation in (30). All irregulars should, thus, be overregularised. However, a more accurate description of the process would be to assume the Variational Model (Yang, 2002), in which different grammars compete in parallel, with the fittest being more likely to be activated. In this view, different events would trigger different competing grammars to enter Stage 2, which means that at any given point, the child could still access the correct listed exception. Since exceptions (irregulars) are attested in the linguistic input, listing of exceptions would over time be favoured. Any hypothesised grammar that does not list exceptions of past-tense formation would quickly fall into oblivion. In our implementation, “competing grammars” should be understood as “competing Vocabulary Items” which all have some probability of being activated during Stage 2. This is one way in which the idea of Vocabulary Items seems to be compatible with learning processes.



**Figure 3:** The three stages of the U-curve illustrated with the corresponding Vocabulary Items. Adapted from Yang (2016).

Secondly, the Tolerance Principle relies on the idea that verb forms are derived from word-formation rules which take the phonological form of a verb as input. In other words, we can say

its input are lexical forms (Biberauer, 2018). For example, the pair *sing-sang* is explained by a rule that takes as input /ɪ/ in the context of /ŋ/ and transforms it into an /æ/, or  $\iota \rightarrow \text{æ} / \_ \eta$ . Whilst we do not currently have an answer as to how the same can be achieved within Distributed Morphology for English irregulars, the same is not true of Brazilian Portuguese, whose verb system and different hypotheses that may be entertained by the child are fully describable in terms of a Vocabulary and Vocabulary Items. This remains an open question to be addressed in future work.

## 5. Final Remarks

This paper aimed to analyse the entire process of overregularisation when children acquire irregular verbs in Portuguese. We showed in section 2 that two previous formal proposals (Lorandi, 2007, 2010; Takahira, 2013) focused only on the stage in which overregularisation is observed; however, they do not fully explain the U-shaped curve, that is, why the child first produces target verbal forms, how these are apparently disrupted, and finally, come back to normality. Another problem with the studies reviewed is that they entail that the child follows a particular principle/operation (Root Fidelity/Minimise Exponence/verb movement) required to produce irregulars and for some unknown reason they stop applying those, producing overregularities, which are later overturned by having that principle or operation return. In a nutshell, these proposals fail either based on empirical evidence or logically by requiring an *ad hoc* account.

A third account reviewed (Wuerges, 2014) has the combined virtues of predicting which verbs may be overregularised in children's speech by looking at the input they received, as well as explaining their non-categorical behaviour when overregularising. Since this account assumed the Variational Model (Yang, 2002), it is compatible with ours, which follows an amended version (Yang, 2016) of that model. Still, this work falls short when having to deal with exceptions, which is why we incorporate its strengths while rejecting some parts of Wuerges' implementation.

Adopting the background of Distributed Morphology and some general learning strategies related to the third factor, such as the Tolerance Principle (Yang, 2016), in section 3, we assigned the overgeneralisation process, when children acquire (ir)regular verbs, to computational efficiency. There is nothing in Universal Grammar that drives the U-shaped curve; that is, there is no failure to follow (innate) principles. In this sense, the overregularisation process results from maximising productive rules. We argued, in section 4, that not having been presented with enough irregularities where  $e > \theta_N$ , little children do not entertain a rule to convey past tense on verbs, which forces them to derive each verb in the past as though it was a singleton (1st Stage), e.g. [ $\sqrt{\text{FAZ}}$ ,  $v$ , PRF.PST, 1, SG]  $\leftrightarrow$  /fiz/ and [ $\sqrt{\text{COM}}$ ,  $v$ , PRF.PST, 1, SG]  $\leftrightarrow$  /komi/, applying Fusion between root,  $v$ , past and AGR in accordance to Siddiqi's (2009) Minimise Exponence principle. Later, as the set of verbs is enlarged, they perceive that the Vocabulary Item /i/ is the productive realisation of the bundle [PRF.PST, 1, SG], given the greater amount of regular verbs

than irregulars,  $e \leq \theta_N$ , and postulate and generalise the Vocabulary Insertion rule  $/i/ \leftrightarrow [\text{PRF.PST}, 1, \text{SG}]$  (2nd Stage). Finally, the child achieves adult knowledge of verb tenses when they realise that the former rule has exceptions, which makes them pair each irregular verb to its target realisation, such as  $[\sqrt{\text{FAZ}}, v, \text{PRF.PST}, 1, \text{SG}] \leftrightarrow /fiz/$ , still maintaining the general/elsewhere rule that  $/i/ \leftrightarrow [\text{PRF.PST}, 1, \text{SG}]$  (3rd Stage).

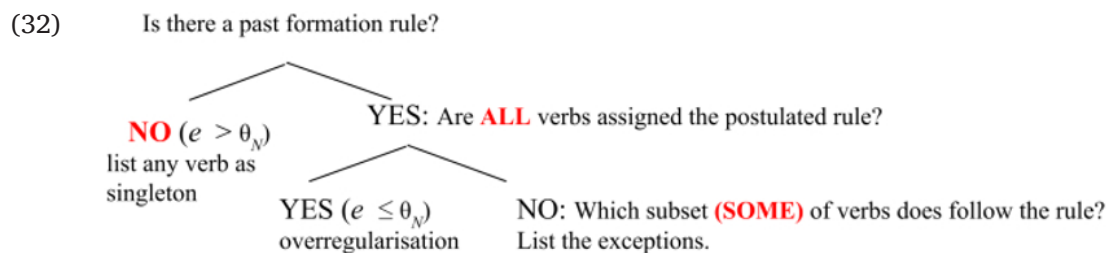
The term “listing” has been used in abundance throughout this work. As we argued in the introduction, we believe that Distributed Morphology seamlessly accommodates for this requirement of the Tolerance Principle, namely that restrictions be listed. For example, List 2, or the list of Vocabulary Items, has a ranking mechanic that makes it instantly compatible with TP. Whenever Vocabulary Items are to be inserted, the one matching the most features in the derivation is chosen. Since the VI for a well-formed irregular verb will always contain its root, it is going to be correctly inserted.

As presented in the foregoing, the path children follow to acquire (ir)regular verbs is very much like the general learning strategy envisaged in terms of Maximise Minimal Means (as proposed in Biberauer, 2011 *et seq.*), in which acquirers do not postulate any specific rule at the beginning, then they postulate a rule and generalise it to all cases, and finally, they perceive the rule only applies to a subset of items. As we claimed, children do not overgeneralise at the beginning because they list Vocabulary Items for each verb, either regular or irregular: both sets of verbs (regulars and irregulars) are part of their Vocabulary, which amounts to saying that the NO rule for past tense formation is still postulated (Stage 1). As a critical mass of VIs is amassed, children have overwhelming evidence that regulars dominate grammar. Therefore, they assign the now productive past formation rule to ALL verbs, hence the overregularisation (Stage 2). As we mentioned before, children do not categorically overregularise during Stage 2, though. This observation can be implemented with the claim that children postulate more than one grammar –  $G_1$  and  $G_2$  –, that is, they have competing grammars, as suggested by Yang (2016). Thus, in  $G_1$ , the child entertains the rule for past formation and follows the ALL stage, overregularising. Additionally, in  $G_2$ , the rule is not generalised, since the rule is applied to only SOME verbs. In this sense, Stage 2 is the transition between ALL and SOME. As argued by Biberauer (2011 *et seq.*), this NO-ALL-SOME path is followed by every child. Hence, children are expected to generalise other instances during language acquisition. Indeed, once they have learnt the word for *dog*, they might generalise it to ALL animals that have four paws and are furry (cf. Oliveira, 1989, p. 49). Furthermore, we also observe the ALL stage regarding morphology itself: children overgeneralise the usage of *des-*, used by adults to convey the reversion of an action, such as in the opposition *fazer* ‘do’/*des-fazer* ‘undo’, *ligar* ‘turn on’/*des-ligar* ‘turn off’. In this way, instead of producing adult forms that are irregular regarding this *des-* pattern, such as *esfriar* ‘cool down’, they produce *desquentar* from *esquentar* ‘heat’ (3;11.10 – cf. Figueira, 2010, p. 122), *desabrir* from *abrir* ‘open’ instead of *fechar* ‘close’ (4;5.17 – cf. Figueira, 2010, p. 140), *deslaçar*



from *laçar* ‘tie’, instead of *tirar o laço* ‘untie’ (4;6.4 – cf. Figueira, 2010, p. 139). In this way, we can liken Stage 2 to children answering ALL to the question of which verbs follow regular rules whenever  $G_1$  is activated; and one could say the U-curve is yet another case of overgeneralising in the grand scheme of language acquisition.

Back to verb acquisition, by the time the child has acquired yet more verbs, and the exceptions – irregular verbs – to the aforementioned rule are perceived in the intake as only in SOME verbs, as it was previously entertained, they only list VIs for individual irregular verbs, as the adult grammar does (Stage 3). Thus, NO > ALL > SOME patterns emerge. Hence, the overregularisation process to acquire (ir)regular verbs ends up being a corollary of computational efficiency in terms of maximising (productive) rules. Our proposal implies that perhaps what guides the child to change from one stage to another, that is, from NO to ALL and finally to SOME, is Yang’s (2016) Tolerance Principle, as in (32). At this point, unfortunately, it is not possible to establish the exact threshold children change from one stage to another. Further research is necessary to investigate whether the threshold for the child acquiring (ir)regular verbs varies cross-linguistically.



## Acknowledgements

We sincerely thank the Journal of Portuguese Linguistics reviewers for their insightful comments and suggestions, which helped to markedly improve the original manuscript. We would also like to thank the mediators and participants at the conferences in which previous versions of this work were presented: namely the 1st CIELIN, the 68th GEL Seminar, and Study Group on Distributed Morphology (GREMD). Finally, Paulo Ângelo de Araújo-Adriano thanks FAPESP (number 2019/17443-9) and Rafael Luis Beraldo thanks CAPES (number 88887.479688/2020-00) for financially supporting this work.

## Competing Interests

The authors have no competing interests to declare.

---

## References

- Anderson, S. R. (1982). Where's Morphology? *Linguistic Inquiry*, 13, 571–612. <http://www.jstor.org/stable/4178297>
- Bassani, I., & Lunguinho, M. V. (2011). Revisitando a flexão verbal do português à luz da morfologia distribuída: um estudo do presente, pretérito imperfeito e pretérito perfeito do indicativo [Revisiting Portuguese Verb Inflection under Distributed Morphology: Realis Imperfect Preterite and Perfect Preterite]. *Revista Virtual de Estudos da Linguagem*, 5, 199–227.
- Bertolino, K. G., & Grolla, E. (2012). O pronome 'ele' está sujeito ao princípio B? Uma discussão sobre resultados experimentais [Is the Pronoun 'He' Subject to Principle B? A Discussion on Experimental Results]. *Revista Linguística*, 8(2), 86–99. DOI: <https://doi.org/10.31513/linguistica.2012.v8n2a4552>
- Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. Cambridge, Mass: MIT Press. DOI: <https://doi.org/10.7551/mitpress/1074.001.0001>
- Biberauer, T. (2011). In defence of lexico-centric parametric variation: two 3rd factor-constrained case studies. Paper presented at the *Workshop on Formal Grammar and Syntactic Variation: Rethinking Parameters*. Madrid.
- Biberauer, T. (2017). Factors 2 and 3: a principled approach. *Cambridge Occasional Papers in Linguistics*, 10, 38–65.
- Biberauer, T. (2018). Less IS More: some thoughts on the Tolerance Principle in the context of the Maximise Minimal Means model. *Cambridge Occasional Papers in Linguistics*, 11, 131–145. DOI: <https://doi.org/10.1075/lab.18080.bib>
- Biberauer, T. (2019a). Children always go beyond the input: The Maximise Minimal Means perspective. *Theoretical Linguistics*, 45(3–4), 211–224. DOI: <https://doi.org/10.1515/tl-2019-0013>
- Biberauer, T. (2019b). Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics*, 45–88. DOI: <https://doi.org/10.5565/rev/catjl.219>

- Bowerman, M. (1982). Evaluating competing linguistic models with language acquisition data: implications of developmental errors with causative verbs. *Quaderni di Semantica*, 3(1), 5–66.
- Câmara, J. M., Jr. (1970). *Estrutura da língua portuguesa* [The Structure of the Portuguese Language]. 14. ed. Petrópolis: Vozes.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton. DOI: <https://doi.org/10.1515/9783112316009>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Massachusetts: The MIT Press. DOI: <https://doi.org/10.21236/AD0616323>
- Chomsky, N. (1969). Linguistics and philosophy. In Hook, S. (Ed.), *Language and philosophy: a symposium*. New York: New York University Press.
- Chomsky, N. (1986). *Knowledge of language: its nature, origin, and use*. New York: Praeger.
- Chomsky, N. (1988). *Language and problems of knowledge: the Managua lectures*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Massachusetts: The MIT Press.
- Chomsky, N. (2001). Derivation by Phase. In Kenstowicz, M. (Ed.), *Ken Hale: A life in language* (p. 27). Cambridge, Mass: MIT Press.
- Chomsky, N. (2005). Three Factors in Language Design. *Linguistic Inquiry*, 36(1), 1–22. DOI: <https://doi.org/10.1162/0024389052993655>
- Cinque, G. (1999). *Adverbs and functional heads*. Oxford; New York: Oxford University Press.
- Cyrino, S., & Lopes, R. (2016). Null objects are ellipsis in Brazilian Portuguese. *The Linguistic Review*, 33(4), 483–502. DOI: <https://doi.org/10.1515/tlr-2016-0012>
- Cyrino, S., & Matos, G. (2002). VP ellipsis in European and Brazilian Portuguese: a comparative analysis. *Journal of Portuguese Linguistics*, 1(2), 177. DOI: <https://doi.org/10.5334/jpl.41>
- Cyrino, S., & Matos, G. (2011). Elipse do VP e variação paramétrica [VP Ellipsis and Parametric Variation]. *Cadernos de Estudos Lingüísticos*, 49(2), 195–206. DOI: <https://doi.org/10.20396/cel.v49i2.8637187>
- Di Sciullo, A.-M., & Williams, E. (1987). *On the definition of word*. Cambridge, Mass: MIT Press.
- Embick, D. (2010). *Localism versus Globalism in Morphology and Phonology*. Cambridge, Mass.: MIT Press. (Linguistic inquiry monograph, v. 60). DOI: <https://doi.org/10.7551/mitpress/9780262014229.001.0001>
- Evers, A., & Van Kampen, J. (2008). Parameter setting and input reduction. In Biberauer, T. (Ed.), *The limits of syntactic variation*. Linguistik aktuell (pp. 483–515). Amsterdam; Philadelphia: John Benjamins Pub. Co. DOI: <https://doi.org/10.1075/la.132.22eve>
- Ferrari-Neto, J., & Lima, M. A. F. de. (2015). Aquisição da morfologia flexional verbal em português brasileiro – um estudo com dados de compreensão [The Acquisition of Verb Inflection in Brazilian Portuguese – A Study with Comprehension Data]. *Revista Prolíngua*, 10(1), 106–120.
- Figueira, R. A. (1996). Os lineamentos das conjugações verbais na fala da criança: multidirecionalidade do erro e heterogeneidade linguística [The Structure of Verbal Conjugations

in Child Speech: Mistake Multidirectionality and Linguistic Heterogeneity]. *Letras de Hoje*, 33(2), 73–80.

Figueira, R. A. (2010). O que a investigação sobre o erro na fala da criança deve a Saussure [What the Investigation of Mistakes in Child Speech Owes to Saussure]. *Cadernos de Estudos Lingüísticos*, 52(1), 115–143. DOI: <https://doi.org/10.20396/cel.v52i1.8637206>

Grolla, E. (2012). Estratégias Infantis na Aquisição da Expressão ‘Ele Mesmo’ em Português Brasileiro [Children’s Strategies in the Acquisition of the Expression ‘Ele Mesmo’ in Brazilian Portuguese]. *Revista Lingüística / Revista do Programa de Pós-Graduação em Linguística*, 8(2), 56–70. DOI: <https://doi.org/10.31513/linguistica.2012.v8n2a4550>

Grolla, E. (2013). A aquisição do Princípio C da teoria de ligação em português brasileiro: questões metodológicas [The Acquisition of Government and Binding’s Principle C in Brazilian Portuguese: Methodological Issues]. *Revista de Estudos da Linguagem*, 21(2), 9–34. DOI: <https://doi.org/10.17851/2237-2083.21.2.9-34>

Halle, M., & Marantz, A. (1993). Distributed Morphology and the pieces of inflection. In Hale, K. & Keyser, S. J. (Eds.), *The view from building 20* (pp. 111–176). Cambridge, MA: The MIT Press.

Halle, M., & Marantz, A. (1994). Some keys of distributed morphology. *MIT Working Papers in Linguistics*, 21, 275–288.

Harris, J. (1999). Nasal Depalatalization no, Morphological Wellformedness sí: The Structure of Spanish Word Classes. In *Papers on Morphology and Syntax, Cycle One* (Vol. 33, pp. 47–82). Cambridge, Massachusetts: MIT Press.

Ippolito, M. (1999). On the past participle morphology in Italian. *MIT Working Papers in Linguistics*, 33, 111–137.

Lopes, R. E. V. (2009). Aspect and the acquisition of null objects in Brazilian Portuguese. In Pires, A. & Rothman, J. (Eds.), *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese*. *Studies on Language Acquisition* (Vol. 35, pp. 105–128). Berlin, New York: Mouton de Gruyter. DOI: <https://doi.org/10.1515/9783110215359.1.105>

Lorandi, A. (2010). Formas morfológicas variantes na aquisição da morfologia: evidências da sensibilidade da criança à gramática da língua [Morphological Variant Forms in the Acquisition of Morphology: Evidence on Child Sensitivity to Language Grammar]. *Letrônica*, 3(1), 81–96.

Maldonado, I. R. (2003). *Erros na aquisição da flexão verbal: uma análise interacionista* [Mistakes in Verb Inflection Acquisition: An Interactionist Analysis] (Unpublished doctoral dissertation). Campinas: Instituto de Estudos da Linguagem, Unicamp.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., John Rosen, T., Xu, F., & Clahsen, H. (1992). Overregularization in Language Acquisition. *Monographs of the Society for Research in Child Development*, 57(4), 1–178. DOI: <https://doi.org/10.2307/1166115>

Oliveira, M. K. de. (1989). Algumas Contribuições da Psicologia Cognitiva [Some Contributions of Cognitive Psychology]. *Idéias*, 6, 47–51.

Oltra-Massuet, M. I. (1999). *On The Notion of Theme Vowel: A New Approach to Catalan Verbal Morphology* (Master of Science in Linguistics), Department of Linguistics and Philosophy, MIT, Massachusetts.

- Pinker, S., & Prince, A. (1988). On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73–193. DOI: [https://doi.org/10.1016/0010-0277\(88\)90032-7](https://doi.org/10.1016/0010-0277(88)90032-7)
- Prince, A., & Smolensky, P. (1993). *Optimality Theory Constraint Interaction in Generative Grammar*. Cornwall: Blackwell Publishing.
- Resende, M. (2021). As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem [Words Babies Don’t Say: Revisiting the Problem of the Idea of “Word” Under Language Acquisition]. *Revista do GEL*, 18(2), 128–159. DOI: <https://doi.org/10.21165/gel.v18i2.3134>
- Roberts, I. (2012). Macroparameters and minimalism A programme for comparative research. In Galves, C. et al. (Eds.), *Parameter theory and linguistic change* (pp. 319–334). Oxford studies in diachronic and historical linguistics. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199659203.003.0017>
- Roberts, I. (2019). *Parameter hierarchies and Universal Grammar*. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oso/9780198804635.001.0001>
- Roberts, I., & Roussou, A. (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge, UK: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486326>
- Roberts, I. G. (2007). *Diachronic syntax*. Oxford; New York: Oxford University Press.
- Santos, A. L., & Lopes, R. E. V. (2017). Primeiros passos na aquisição da sintaxe: direcionalidade, movimento do verbo e flexão [First Steps in Acquiring Syntax: Directionality, Verb Movement and Inflection]. In Santos, A. L. & Freitas, M. J. (Eds.), *Aquisição de língua materna e não materna: questões gerais e dados do português* [First and Second Language Acquisition: General Issues and Portuguese Data] (pp. 155–175). Berlin: Language Science Press, 2017. DOI: <https://doi.org/10.5281/zenodo.889429>
- Siddiqi, D. (2009). *Syntax within the word: economy, allomorphy, and argument selection in Distributed Morphology*. Amsterdam; Philadelphia: John Benjamins Pub. Co. DOI: <https://doi.org/10.1075/la.138>
- Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. *Journal of Verbal Learning and Verbal Behavior*, 18(5), 523–534. DOI: [https://doi.org/10.1016/S0022-5371\(79\)90284-6](https://doi.org/10.1016/S0022-5371(79)90284-6)
- Takahira, A. G. R. (2013). O processo de aquisição de verbos irregulares no português brasileiro [The Irregular Verb Acquisition Process in Brazilian Portuguese]. *Estudos Linguísticos*, 42(1), 430–441.
- Tomasello, M. (2005). *Constructing a language: a usage-based theory of language acquisition*. Cambridge: Harvard University Press. DOI: <https://doi.org/10.2307/j.ctv26070v8>
- Wuerges, T. E. (2014). *A aquisição da morfologia verbal por crianças falantes de português brasileiro e o uso de formas variantes* [The Acquisition of Verbal Morphology by Brazilian-Portuguese-Speaking Children and the Use of Variant Forms] (Unpublished masters’ thesis). São Paulo: Universidade de São Paulo, 2014.
- Yang, C. D. (2002). *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yang, C. D. (2016). *The price of linguistic productivity: how children learn to break the rules of language*. Cambridge, Massachusetts: The MIT Press. DOI: <https://doi.org/10.7551/mitpress/9780262035323.001.0001>

Yang, C., & Roeper, T. (2011). *Minimalism and language acquisition*. In Boeckx, C (org.) *The Oxford handbook of linguistic minimalism* (pp. 551–573). Oxford handbooks in linguistics. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780199549368.013.0024>

Yip, K., & Sussman, G. J. (1997). Sparse representations for fast, one-shot learning. *AAAI-97 Proceedings*.

