Isogloss
Open Journal of Romance Linguistics

Received: 30-05-2024
Accepted: 27-09-2024
Published: 02-10-2024

What lexical acquisition has to say about a non-lexicalist architecture of grammar – and vice-versa

Rafael Luis Beraldo
Universidade Estadual de Campinas (Unicamp)
rafaelluisberaldo@gmail.com

Paulo Ângelo Araújo-Adriano
Universidade de São Paulo (USP)
pauloangelo@usp.br

How to cite: Beraldo, Rafael Luis & Paulo Ângelo Araújo-Adriano. 2024. What lexical acquisition has to say about a non-lexicalist architecture of grammar – and vice-versa. In *30 Years Distributing Morphology from North to South: the view from Romance*, eds Maurício Resende & Ana Paula Scher. Special issue of *Isogloss. Open Journal of Romance Linguistics* 10(6)/6, 1-33.
DOI: https://doi.org/10.5565/rev/isogloss.460

Abstract

Distributed Morphology (DM) predicts that the units of syntactic derivation are smaller than words. This paper explores the implications of this prediction for language acquisition research and questions if DM is descriptively and theoretically sound when faced with acquisition phenomena. We first introduce independent evidence supporting sub-word units in acquisition: results from a computational model of lexical acquisition show that slightly more morphologically complex input data, such as Brazilian Portuguese when compared to English, cause a substantial decrease in the model's performance; children's early productions when acquiring polysynthetic languages reveal they are attempting to find these languages' morphological units, instead of relying on chunks of non-analysed material; and words are shown to lack explanatory power in describing language acquisition in terms of storage, bootstrapping, or production. We then bridge the gap between DM and earlier proposals for the identification of words and formal features, briefly outlining a strategy for acquiring morphemes. Finally, we present accounts for two prevalent phenomena linked to language acquisition through the lens of DM: the overregularisation in acquiring irregular verbs, a step observed in children acquiring different languages; and the Brazilian Portuguese verbal paradigm shift, an example of morphological diachronic change. Our findings support the non-lexicalist derivation of words, highlighting that words are incompatible with acquisition from multiple perspectives. This suggests that understanding language acquisition benefits from considering the smaller, morpheme-based units predicted by Distributed Morphology.

Keywords: lexical acquisition, Distributed Morphology, overregularisation, diachronic change.

1. Introduction

Since Lees' (1957) review of Chomsky's Syntactic Structure, language acquisition has become the mainstay for validating theories of grammar. Assuming that Distributed Morphology (DM; Halle & Marantz 1993) proved profitable for the architecture of grammar, after 30 years of distributing morphology from south to north[1], it becomes legitimate to ask what the direct consequences of a non-lexicalist model are to language acquisition. The central question we address in this paper is: Can DM be descriptively and theoretically sound regarding the unit of acquisition, which is not made up of words but of smaller units?

We start by reviewing independent arguments that favour units smaller than words in lexical acquisition (section 2). From the standpoint of learnability, an important factor in evaluating formal theories of language, we show that a computational model grounded in lexicalist approaches struggles with small increases in morphological complexity due to its input being switched from an English *corpus* to a Brazilian Portuguese (BP) *corpus* (Faria 2005). Then, we present compelling examples of early morpheme production from the acquisition of polysynthetic languages, in which an utterance and a morphological word are on equal terms (Kelly *et al.* 2014; Allen 2017). We also review arguments supporting the non-special status words have in languages in which they seem to exist (Resende 2021a). These considerations hint at the fact that children do not seem to learn words, but rather morphemes. Drawing from DM and previous works on lexical acquisition (Christophe *et al.* 2008; Corrêa 2009), we advance a morpheme-based lexical acquisition hypothesis according to which children are always looking for the smallest units available in the input data. By following this principle, we claim that computational models could be made more resilient to data with varying morphological complexity; children, in turn, could be described as learning languages with differing morphological complexity in a similar manner, being able to identify the roots, affixes, and features that compose the words they hear.

In what follows, we present a theoretical treatment, based on DM, of two specific aspects of language acquisition. First, the overregularisation in acquiring irregular verbs, giving rise to the known U-shaped curve (Marcus *et al.* 1992), is explained in terms of DM's lexical architecture (section 3). Then, assuming that there is a crucial link between language acquisition and diachronic change (Lightfoot 1979),

[1] We would like to express our gratitude to the organisers and attendees of the event "30 Years Distributing Morphology from North to South (DM30)," held from October 3–6, 2023, at Universidade de São Paulo (USP), for their invaluable feedback on this work.

we provide an account for the verbal paradigm change in BP from six forms to only two based on DM's notion of Vocabulary items (section 4).

All in all, we show that words are not compatible with acquisition whether we look at computational models, at the mounting empirical evidence from complex morphology languages, at theoretical arguments from syntacticians, as well as at the very architecture of grammar posited by DM. Our proposal predicts both the overregularisation of irregular forms, caused by a rearranging morphological inventory, as well as linguistic change, where underspecifications in pairs of phonological exponents and features could give rise to a different system in a certain span of time. This demonstrates that diachronic change and language acquisition can indeed be analysed in complementary ways. In line with Distributed Morphology, we argue that the non-lexicalist derivation of words has positive descriptive consequences for language acquisition.

2. Learnability, Language Acquisition, and Grammar Theory

Language acquisition and grammar theory have long been intertwined, forming a research feedback loop that shapes the learning mechanisms of acquisition as well as the different models of language proposed. This interaction poses two questions, which at the same time exist on their own, as do the two sides of a coin, but cannot be answered in complete isolation. The first question is whether a given language design is compatible with language acquisition or not. This played no small part, as we shall review, in the different stages of Generative linguistics. The second question asks whether the resulting representations of linguistic knowledge are compatible with a viable learning procedure, *i.e.* one that children can be reasonably expected to employ, and has largely been identified with the field of learnability. We now turn to a more detailed examination of the answers given to these two questions.

2.1. Language Acquisition, and Grammar Theory

Although the term acquisition never appeared in Chomsky (1957), it was from Lees' (1957) review of that book that a handful of works in grammar theory started to recognise that language acquisition could be the linchpin for testing models of grammar. Lees notes that if it was possible to align linguistic theory and language acquisition, one has the upper hand. Since then, the question of how children rapidly acquire language despite limited and often imperfect linguistic data has motivated different linguistic agendas. This impasse, known as Plato's Problem, found a promising solution in Chomsky's (1986, 1988) proposal of a dedicated language acquisition device that is biologically determined and hence innate, known as the Faculty of Language (FL). From it, guiding the specific properties of a given language, general principles and parameters would comprise Universal Grammar (UG), the initial state of FL. Linguistic stimuli, the input, feed UG, shaping how parameters will be set and, eventually, which language will be acquired.

In light of Plato's Problem, Principle and Parameter (P&P) theory (Chomsky 1986) offered a compelling explanation for children's rapid acquisition of language. By positing a set of innate principles (the UG), P&P theory posited that language acquisition primarily involves exposure to linguistic input, followed by a setting of parameters to values specific to the target language. Hereby, as Yang and Roeper (2011) point out, parameters were a manner to simplify the task of acquiring a language. With the emergence of the Minimalist Program (Chomsky 1993), FL starts to be taken as a broad perceptual and cognitive system, not attributing the totality of linguistic properties to UG but including a third factor, principles not specific to FL, such as principles of efficient computation (see Yang 2002).

Considering learnability and language acquisition under Minimalist grounds, harkening back to the beginning of the Generative Grammar enterprise, Yang & Roeper (2011: 560) set forth three investigative axes, one of which is of interest to us: How would a theory of grammar simplify the learner's task to achieve successful acquisition with a relatively small amount of data? In the feedback loop of linguistic research, if language acquisition informs the current model of language, how well do the cognitive faculties, domain-specific or not, guide children's hypotheses of the properties and structure of what is to be learned?

2.2. Language Theory and Learnability

The view that children must have innate knowledge about the languages they are to acquire can be traced back to a problem of learnability. After Gold's (1967) demonstration that an underlying generative mechanism for the languages in the Chomskian hierarchy could not be identified without negative evidence, language acquisition researchers proposed that a theoretical learner, in order to be successful, would have to deal with a reduced hypothesis space (Wexler & Culicover 1983). The determination of which faculty-specific expectations children bring to the task later led to the development of P&P theory, part and parcel of doing Generative linguistics. In the forthcoming, we introduce computational models of lexical acquisition to show how a model (Faria 2015), due to its input data being represented as morphological words, fails to perform when input is changed from English to BP, as well as how non-lexicalist theories of grammar could provide an answer to this issue of learnability.

Arguments based on learnability depend on the accuracy of our theories of language and of learning procedures: "The problem is to find out exactly what are possible human languages and how these are learned." (Wexler & Culicover 1983: 2) Determining the available data, its shape and features, is also fundamental (Atkinson 1997). Whenever problems arise, the solution lies in tweaking one or all of these variables (Atkinson 1997): The assumptions about the mental representation, about the data, or about the learning procedure are investigated to determine which is wrong. Computational models are particularly well-suited for investigating this last aspect, since the same dataset can be fed into different algorithms in order to determine which algorithm produces the expected mental representations (Pearl 2010).

Computational models of lexical acquisition compete to show which learning procedure[2] is able to extract a lexicon out of input data, generally consisting of pairs of utterances and conceptual representations, in a way that is compatible with child development (Siskind 1996), with the probabilistic nature of the mind (Fazly, Alishahi & Stevenson 2010), or with the reasonable assumption that children have limited memory (Trueswell *et al.* 2013; Yang 2019). All of the cited models try to solve the

[2] The terms 'acquisition' and 'learning' are used interchangeably in the context of computational models. Similarly, 'word' refers to a 'morphological word,' unless otherwise specified.

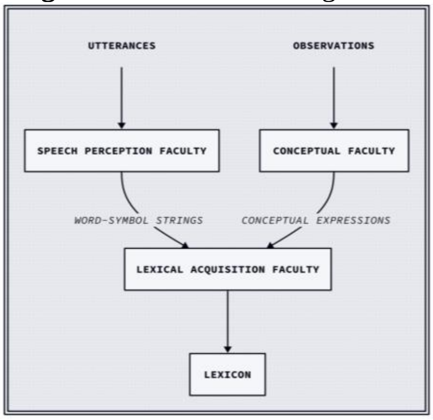
problem of referential uncertainty (or *gavagai* problem; Quine 1960). In short, this problem, a staple issue in lexical acquisition, asks the question of how one can correctly infer what an utterance in an unknown language is about, given that extra-linguistic context can be construed in virtually infinite ways through language. Hearing words they could not have any prior knowledge about, a newborn child is faced with just this question. For example, the word *toast*, spoken in a sentence such as "What did you think of the toast?" at the breakfast table, could refer to anything from the TABLE to the DOG sniffing around it.[3]

The solution and unifying idea behind these computational models – the cross-situational strategy – is that a learner could stay relatively neutral about which word-meaning associations are correct in earlier stages, storing meaning hypotheses for later checking. As time goes by, meaning consistent across all (or most) contexts would prevail, whilst inconsistent (and thus wrong) meaning hypotheses would be discarded. First proposed in theoretical works on lexical acquisition (Pinker 1989; Fisher *et al.* 1994; Gleitman 1990), the cross-situational strategy has been more thoroughly specified in algorithmic form. Most of these works implement an online, probabilistic learner, that is, one that updates its knowledge with each new input datum, rather than by processing the *corpus* all at once, producing a distribution of association probabilities between words and meanings (Beraldo 2020).

Although these models vary in the particular learning procedure advanced (for more recent insights, refer to Stevens *et al.* 2017; and Yang 2019), they tacitly assign a special treatment to words in the learning procedure. The prevailing view, elucidated notably in Siskind (1996), asserts that utterances are split into "word-symbol strings," generated by a speech perception faculty, whereas a conceptual faculty generates "conceptual expressions" based on real-world "observations". A lexical acquisition faculty then takes these two constructs and outputs a lexicon (see Figure 1). The primary focus is on which algorithm extracts a lexicon in alignment with child development factors while being conservative in memory requirements. However, we will direct our attention towards the assumptions about the formal structure of the "word-symbol strings," as well as their implications for learnability.

[3] The problem is further aggravated when considering all five difficulties listed by Gleitman (1990) and reviewed by Siskind (1996): (1) utterances might have multiple words, which increases indeterminacy; (2) more than a single meaning hypothesis is often compatible with any given conversational setting; (3) related to that, all the hypotheses a child might conceive of could be wrong, leading to spurious word-meaning associations; (4) children start with no known words, having to bootstrap lexical learning seemingly out of thin air; and (5) there are homonyms, or associations from one sound to multiple meanings, such that more than one lexical entry might be afforded in some cases, as is the case with *bat*, both an animal and a club used in some sports, such as baseball.

Figure 1. Lexical Learning Model



Source: Adapted from Siskind (1996: 44)

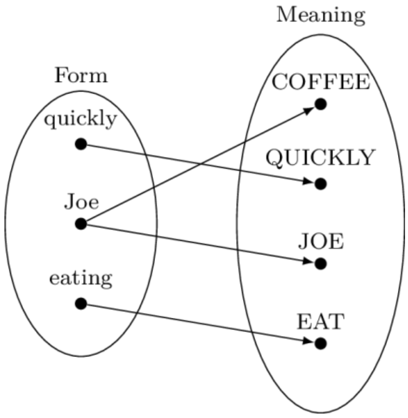
This model of lexical learning tacitly assumes that "words", that is, morphological words, are the basic formal unit in lexical acquisition. In Atkinson's (1997) terms, the data, as well as the mental representation, are taken to be appropriately represented by the morphological word. Undoubtedly, this view is based on previous theoretical studies on lexical acquisition, which normally avoid the age-old issue of the prevalence of the word as an actual linguistic unit. When comparing the definitions for *word* given in the literature on lexical acquisition, different and somewhat incompatible ones are found: Bloom (2000) defends *Saussurean signs* are the unit of lexical acquisition, whereas Golinkoff & Hirsh-Pasek (2000) state that it is the *minimal free form*. Throughout this paper, we will focus on the issues that might arise from taking the word as a starting point.

First, consider how a computational model simulates the task of lexical acquisition, through a prototypical example offered by Fazly, Alishahi & Stevenson (2010). Utterances from CHILDES *corpora* (MacWhinney 2014) are paired with meaning representations in order to simulate the two faculties posited above:

(1) Utterance "Joe is quickly eating an apple."
Conceptual Representation {JOE, QUICKLY, EAT, A, BIG, RED, HAND}

The model is tasked with generating a lexicon, that is, a mapping between each word and its meaning, *e.g.* {*Joe* → JOE, eating → EAT, …} (see Figure 2). As the model encounters more uses of each word, its meaning representation is gradually refined. The example above captures some of the challenges for lexical acquisition in Gleitman (1990). Referential uncertainty is simulated by the inclusion of confounding meanings, such as RED (presumably from the mention of an apple) and HAND (from the people present). The example is also noisy, since APPLE is missing from the meaning hypotheses the learner postulated, which would lead to a wrong hypothesis as to what *apple* means. Finally, *Joe* can sometimes mean COFFEE, as in "a cup of joe," which is a case of homonymy, which some models account for.

Figure 2. Form-meaning mappings



Source: Own work

Now, consider how the speech perception faculty is represented by this model, which accurately exemplifies most, if not all, cross-situational models of lexical acquisition. It is assumed that the learner is able to extract perfect, dictionary-like morphological words from speech. This is, of course, a best-case scenario simplification, since the main goal of such models is to attest to the viability of a given word-learning algorithm.

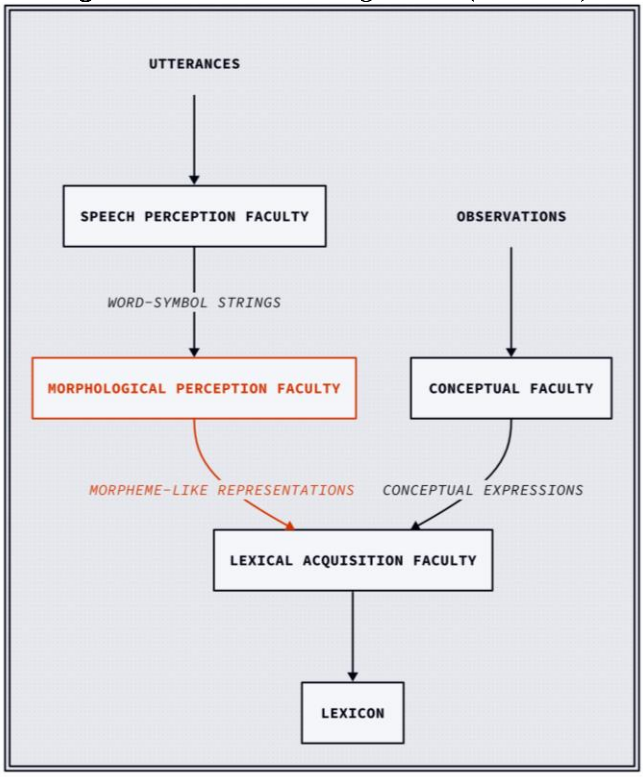
One would not expect, however, that changes in the input data, the "utterances", and thus in the output of the speech perception faculty, would have an impact in the performance of the learning algorithm implemented by any given model. In other words, if the learning mechanism is language-agnostic, being able to extract a lexicon from any given *corpora* of any language, as long as the model is given a set of words paired with a set of meaning hypotheses. There is evidence to the contrary, showing that changing a model's input from English to BP negatively impacts its performance. This poses a problem of learnability in the sense of Atkinson (1997: 91).

Such a problem of learnability is found in Faria (2015), a cross-situational study of lexical acquisition for English and BP. Faria's model, a re-implementation of Siskind's (1996), tested three *corpora* with differing results. The model was able to converge on 95% of words[4] when the input data came from an English *corpus*, whereas only 39% to 53% of words correctly converged for BP *corpora*. The author speculates (Faria 2015: 8) that, since the performance for English data was similar to Siskind's original results, the culprit might be the higher lexical sparsity in BP data. This lexical sparsity emerges, for example, from the higher number of inflected forms any given verb might have in BP *versus* in English. Take, for instance, the verb *trabalhar* ('to work'), which could be manifested in as many as 42 forms (*e.g., trabalha* 'work.PRS.3.SG', *trabalhamos* 'work.PRS.1.SG', *trabalharia* 'work.COND.3.SG', *trabalhando* 'work.GER', *etc*.), whereas its English counterpart would have only four (*work, works, worked, working*). Unaware of the underlying lexeme, the model disperses its representation of the BP verb over its many inflected forms, consequently having less chances to converge on any given verb. The same should hold for words

4 That is, there were correct lexical entries for 95% of the words in the input.

such as nouns and adjectives, since these classes also have more inflections in BP than in English.

Figure 3. Lexical Learning Model (amended)



Source: Adapted from Siskind (1996: 44)

From the standpoint of the learnability of word-meaning pairs, it would seem that an interesting challenge is posed by languages more morphologically complex than English. One solution, given by Faria (2015) himself, would be

> […] decomposing words into stems and affixes, which by hypothesis could eliminate the problem of sparsity both by guaranteeing frequent expositions to the stems and by assigning affixes to the category of functional words. (Faria 2015: 8)

Indeed, this would make the BP input utterances more similar to (1), that is, more English-like. We argue that lexical acquisition has been traditionally characterised as the learning of mappings from morphological words onto meanings (Gleitman 1990; Markman 1990, 1994; Fisher *et al.* 1994; Golinkoff & Hirsh-Pasek 2000; Bloom 2000; a.o.) only because English words, and particularly nouns, are so seemingly atomic in their form. These theories then inform computational models, which turn to data from *corpora* of English-speaking children and caretakers.

No claim is made that the core acquisition mechanisms in these theories and models are wrong by any means. A simple lemmatiser could be added after the speech perception faculty (see Figure 3) and this step would probably bring Faria's (2015) model back in line with expected performance. Rather, the argument being made is that there are grounds to review the mental representations assumed in these models of learnability (Atkins 1997).

We argue that the shortcomings in learnability found in Faria (2015) hint at the fact that acquiring a lexicon does not amount to storing mappings between morphological words and meanings. In section 3.1, we follow Faria's suggestion that the input data should be analysed in terms of smaller elements, or *morphemes*, which is aligned with the assumptions of DM. Before that, however, we will explore more arguments against words as special units of acquisition, as well as a non-lexicalist model of grammar that offers alternative tools to explain lexical acquisition.

2.3. More Arguments against Word-based Lexical Acquisition

Faria's (2015) computational model suggests that when a mechanism of lexical acquisition is actually implemented, problems arise if the analysis of linguistic input data does not go further than the morphological word. In fact, it would seem that the more complex the morphology of the language used as input, the lower the rate of convergence that should be expected from such computational models. This stems from the fact that in English – virtually the only language informing theories of lexical acquisition – many "content words", often the elected target of lexical acquisition, coincide with their roots. Thus, the assumption that the task of lexical acquisition is sufficiently modelled as a mapping between a word token and a meaning token, such as *jar* → JAR, does not hold for a language like BP, where *jarro* and *jarra* (respectively 'clay jar' and 'glass jar'; the theme vowels *-o* and *-a* do not contribute to the meaning of these words) compete with each other to form a mapping to the concept JAR (see Resende & Santana 2019: 28).

In a review of studies of the acquisition of polysynthetic languages, Kelly *et al.* (2014) also point out the shortcomings of basing whole theories of lexical acquisition on a single language. One particular contrast is between languages like English – which the authors classify as being more *isolating* – and polysynthetic languages, in which a "single verb can express what would in English take a multi-word clause" (Kelly *et al.* 2014: 52). A particularly extreme example in (2) comes from Amuesha, an Arawak language (Aikhenvald 2012: 129):

(2) Amuesha, Aikhenvald (2012: 129):
Ø-omaz-amy-eʔt-ampy -es-y-e.s-n-e.n-a
3SG-go.downriver-DISTRIBUTIVE-EPENTHETIC-DATIVE.ADVANCEMENT-EPENTHETIC-PL-EPENTHETIC-late-PROGRESSIVE-REFL
'They are going downriver by canoe in the late afternoon stopping often along the way'

This example, which is at the same time an utterance and a morphological word, raises questions as to how a cross-situational, mapping-based learner would operate, as well as to the nature of making word-meaning associations, illustrating the need for cross-linguistic investigation of theories of lexical acquisition.

Besides the obviously naive strategy of mapping a whole complex word into a whole complex utterance, two more reasonable strategies seem to be available to the child: Either segmenting the input data into unanalysed "chunks", which are then mapped into rudimentary meanings (Brown 1973), or looking for the smallest units of meaning and mapping these into rudimentary meanings. The studies reviewed by Kelly *et al.* (2014) indicate children acquiring polysynthetic languages favour the latter approach. When morphemes are more transparently placed (*i.e.* more phonetically salient), early productions reveal children successfully extract bare roots from the input data. This is the case of a child (1;1) acquiring Navajo who produced *da* ('sit'), based on the adult form *ní-d'aah* (THEME/2SUB-sit; Kelly *et al.* 2014: 56). In more opaque languages, in which the root is obscured by surrounding phones, children's first productions are the stressed segments in phonological words. For example, a child (1;9) acquiring Mohawk produced *'ti* in place of the adult form *sa'tita* ('get in'; Kelly *et al.* 2014: 54). The authors show that the early production of bare roots or of more salient segments is consistent across polysynthetic languages.

Based on this data, Kelly *et al.* (2014: 61) dismiss the hypothesis that non-analysed chunks of data inform early lexical acquisition, but rather argue that children seek to extract material from within morphological words, which are often highly templatic in nature in polysynthetic languages. The same conclusion is arrived at by Allen (2017), who describes the acquisition of Inuit in much the same way. In this language, children's first productions are monomorphemic, despite the input words varying between two and ten morphemes in length. The author classifies children's productions into three types: Particles, which are free forms (*e.g., Auka*, 'no'), isolated nouns (*Piipi*, 'baby'), and unflexed roots (*Apaapa*, 'food, eat'). Since these roots are accompanied almost exclusively by affixes, Allen argues (2017: 8) that children could not have memorised non-analysed forms. Although relatively few studies investigate the acquisition of languages with highly complex morphology, it is evident that the morphological word cannot be the unit children rely on when learning a lexicon. Several other authors have further developed this point.

Resende (2021a) more incisively negates the special status of the word in language acquisition in general. His tripartite argument relies on acquisitional facts to state that the child does not store words in their mental lexicon, does not prioritise the (morphological) word to acquire lexical material, nor produces words independently at any moment, consequently rejecting the so-called holophrastic period. In the following, we briefly consider the evidence.

First, Resende revisits Marantz's (1997) arguments that the word cannot be a "special unit" in language with a different dataset, once again concluding that the word is not special for storing either phonological material or meanings. This is because the prosodic word can be larger or smaller than the morphological word (Resende 2021a: 5): *guarda-chuva* ('umbrella') is a 2:1 word (two prosodic units to one morphological word: ['guar.da.'ʃu.va]), whereas *o menino* ('the boy') is a 1:2 word ([o.me.'ni.no]). Analogous to what we saw above, morphemes, phrases, and even sentences can have idiomatic meanings (Resende 2021a: 6), *i.e.*, secondary meanings that overlap with those normally associated with a root, phrase, or sentence. This is the case, for example, with *pianinho* (lit. 'small piano', meaning calm, discreet), where the root √PIAN and the diminutive suffix *-inh-* do not contribute their canonical meanings. Thus, the word is an ineffective unit for analysis and storage: The child must constantly look inward and outward for truly relevant phonological and semantic units.

Next, the author reviews a series of psycholinguistic studies investigating children's perception of morphological and syntactic phenomena. An example is the distinction between nouns and adjectives examined by Teixeira & Corrêa (2008). Employing pseudowords as experimental stimuli, the researchers demonstrate that infants aged 18–22 months analyse structures containing determiners, nouns, and adjectives based on their distribution alongside considering adjectival suffixes. Thus, the child classifies the pseudoword *dabo* as N in *"um dabo mipo"* (along the lines of "a mipo dabo"), but *daboso* as ADJ in *"um daboso mipo"* ("a mipo dabous"), notwithstanding this being the non-canonical word order in BP, arguably because of the suffix *-oso*, associated with adjectives. Additionally, Resende references further studies (*cf.* Name 2008; Bagetti & Corrêa 2010; Takahira 2013; Ferrari-Neto & Lima 2015; a.o.) that adopt diverse experimental methodologies, revealing that children typically engage in analyses either more or less granular than the word level. The only instance of apparent coincidence is attested by Albuquerque, Bezerra & Ferrari-Neto (2012), who study the perception of the productive reversal prefix *des-* (similar to *un* in 'undo') in 54-month-old children.[5]

To conclude, Resende (2021a) revisits the technical discussion advocated by the proponents of Distributed Morphology (DM). He argues that given the absence of unanalysable forms in the input available to children and their early morphological abilities, it is reasonable to assume that children's very first productions are already the product of a linguistic engine that concatenates morphemes. Even the seemingly primitive word *bola* ('ball') would be the outcome of a derivation that, at the very least, consists of the root √BOL and the theme vowel *-a*. Although early productions are describable in terms of holophrases, isolated words, or even telegraphic speech, these are not unanalysable forms. Instead, from the outset, children's utterances are grounded in syntactic structures, however minimal they may be.

In the foregoing, we showed that Faria (2015) suggested that a computational model of lexical acquisition, in order to be cross-linguistic, has to include morphological pre-processing. Further, Kelly *et al.* (2014) and Allen (2017) argued that children learning polysynthetic languages are aware of morphemes from early on. Finally, Resende (2021a) theoretically showed that the word is not an adequate unit for storing, bootstrapping, or analysing children's early linguistic productions. Based on these considerations, in the following, we present a model for the architecture of language that treats words without any special status.

2.4. Distributed Morphology as a well-suited model for explaining lexical acquisition

Hence, as we saw, mental representations cannot be based upon words; which amounts to saying that one needs an alternative model for the acquisitional procedure. Distributed Morphology, being a non-lexicalist model of language, presents a compelling set of features that might prove useful in better characterising what the acquisitional input data actually looks like.

Embedded in the theoretical framework of the Minimalist Program – MP – (Chomsky 1993), the operation of Merge also serves as the fundamental syntactic

[5] The authors observe, in a preferential looking experiment, that children perform morphological decomposition in the case of derived verbs (*desamassar*, lit. 'uncrumple'), but do not seem to decompose so-called ambiguous verbs, such as *desmanchar* (never meaning 'unstain', but always 'dissolve', 'break apart', 'become undone') and *descobrir* (never meaning 'uncover', as in remove the cover from, but always 'discover'). For example, say the stimuli in a given experimental round are *desmanchar* and two images, one representing the action of unstitching something and the other, of removing a stain from something. The authors expected that if the child performed morphological decomposition, they would preferably look at the second image, which was not verified. However, it is possible that children always opt for morphological analysis, which, if unsuccessful, leads them to less granular levels of analysis, causing the child to determine the prevailing meaning of *desmanchar*. Thus, Resende (2021a: 9) argues that the bootstrapping of lexical acquisition occurs by means of the minimal unit (the morpheme), as opposed to the word.

mechanism. Expanding on this, DM posits that its primitives are subject to the same mechanisms, restrictions, and operations of MP. This logical connection positions Distributed Morphology as a model of syntax applied to words: Just as sentences are to syntax in MP, words are to syntax in DM.

In the Distributed Morphology approach, syntax is the only generative component of grammar, which takes words not as a standalone entity but as a product of various syntactic, semantic, and phonological features distributed across three lists, accessed at different moments of the derivation. List 1, Strict Lexicon, contains features manipulated by syntax. List 2 contains pairings of phonological features to morphosyntactic features (Vocabulary Insertion). And List 3, or Encyclopaedia, interprets the structure based on specific contextual instructions. This distributed nature of word formation significantly impacts our understanding of language structure and acquisition.

List 1 supplies syntax with abstract morphosyntactic and semantic features, lacking phonological (*cf.* Halle & Marantz 1994) and non-compositional semantic content. This initial list equips the syntactic computational system with the necessary elements to construct structures, *viz.*, an inventory of (bundle of) features, roots, and categorisers, generating not only phrases and sentences but also words (Marantz 1997). For instance, to generate a word such as *flor* 'flower', the root $\sqrt{\text{FLOR}}_{[+\beta]}$, a nominaliser feminine $n_{[\text{FEM}]}$, are the elements from Strict Lexicon that will feed syntax from the N(umeration) in (3).[6] In the syntactic component, $\sqrt{\text{FLOR}}_{[+\beta]}$ and $n_{[\text{FEM}]}$ Merge. The structure generated is sent to Phonological Form (PF) and Logical Form (LF).

(3) List 1
N: $\{\sqrt{\text{FLOR}}_{[+\beta]}, n_{[\text{FEM}]}\}$

In PF, triggered by morphological well-formedness conditions, the structure generated can receive additional operations by the Morphological Structure, such as Fusion, Fission, Impoverishment, and node Insertion. In the case of *flor*, a thematic node hosting a nominal thematic vowel is inserted in the *n* categoriser (*cf.* Harris 1999). After these operations after syntax have taken place, the structure is ready to receive phonological content. In List 2, phonological features will be paired with morphosyntactic features, known as Vocabulary Insertion. In the case of *flor*, in Vocabulary, it is defined that for the [+β] class feature given birth to by the root, the ⃠ theme vowel will be inserted, (*cf.* (4a)), along with the pairing of the root $\sqrt{\text{FLOR}}$ with its phonological content /'floɾ/ (*cf.* (4b)).

(4) List 2
a. ⃠ ↔ [+β]
b. /'floɾ/ ↔ $\sqrt{\text{FLOR}}$

Simultaneously, in Logical Form, the Encyclopaedia pairs non-compositional semantic content with the contextual structure generated by syntax. For instance, the instructions in List 3 to interpret *flor* as a noun consider that $\sqrt{\text{FLOR}}$ was merged with a nominaliser; on the contrary, if the root $\sqrt{\text{FLOR}}$ had merged with a verbaliser (and

6 Along with Harris (1999), Alcântara (2003, 2010), and Resende & Santana (2019), we assume that the information about classes is inherent to roots (but see Coelho & Araújo-Adriano 2021 for the contrary) and that invariant non-interpretable gender is a property of the nominaliser (Kučerová 2018; Resende & Santana 2019).

with some abstract temporal morpheme) in syntax, the interpretation will have a different entry to determine the root in this specific context, as sketched below in (5).

(5) List 3
a. √FLOR ↔ "part of a plant that is often coloured" / [[__]√ $n^{o}_{[FEM]}$]
b. √FLOR ↔ "to produce flowers" / [[__]√ v^{o}]

As evident, the concept of a word is distributed across multiple lists within the grammar's structure rather than being centralised in a single location, specifically the Lexicon.

To support the absence of a generative lexicon in DM, Marantz (1997) argues that there are no phonological or semantic phenomena restricted to the domain of the word. According to the author, the notion that phonological operations are limited to the word is not observed in practice, as there are, for example, operations of sound material change conditioned by the type of verb complement. Furthermore, it is not possible to say that the word is inherently the minimal unit of non-compositional meaning, since languages display plenty of idiomatic expressions composed of several words, such as *chutar o pau da barraca* (lit. 'kick the stick of the tent', or to act without care for consequences) and *comer barriga* (lit. 'eat belly', similar to 'drop the ball'). Marantz contends that there are syntactic restrictions, but never lexical ones, on the size of such expressions. In general, for syntactic reasons whose explanation goes beyond our focus, the hard core of idiomatic expressions is configured in the phrase formed by the verb and its internal argument. Therefore, we can modify the subject and say *x chutou o pau da barraca*, where *x* is any DP. Figueiredo Silva & Medeiros (2016) point out that, in fact, this subject flexibility is possible even for sentential expressions such as *a vaca foi pro brejo* (lit. 'the cow went to the swamp', meaning a situation went awry), since *"com essa chuva, nosso passeio foi pro brejo"* ('with this rain, our trip went to the swamp'; Figueiredo Silva & Medeiros 2016: 127) is valid. Thus, the notion of the word as either a morphological or a lexical primitive, from both a phonological and semantic perspective, must be abandoned in the face of data.

In this sense, considering that language acquisition serves as methodological support guiding linguistic theories, in light of Lees (1957: 408) who asserts that "the simplest model we can construct for this [acquisition] reveals that grammar is of the same order as a predictive theory," linguistic theory is only adequate if it can explain language acquisition[7]. Under such circumstances, the question that naturally arises is whether a model that does not prioritise words, such as DM, can also account for language acquisition. Then, if DM is on the right track, the working hypothesis is that it would also explain language acquisition in many aspects. From now on, we entertain this idea, discussing how a non-lexicalist theory, here represented by DM, can deal with language acquisition.

[7] An anonymous reviewer questioned the psycholinguistic reality of Distributed Morphology List 3. We thank the comment, and we agree that List 3, prima facie, might pose a problem from the psycholinguistic point of view, for it combines both encyclopaedic knowledge and "lexical" semantics (see for instance Minussi & Bassani 2021 and Resende 2020, for whom lexical semantics is in the root itself). Going beyond our proposal in this paper, one should face the DM approach with psycholinguistics reality to test whether DM achieves what Guimarães (2017) calls neurophysiological adequacy; after all, finding some sort of psychological reality to a given theoretical approach is the desideratum of any grammar model, since the early stages of Generative Grammar.

3. A Morpheme-based Hypothesis for Lexical Acquisition

In order to fully realise the idea that linguistic theories must be compatible with acquisitional realities, we must first specify how a non-lexicalist theory of "lexical" acquisition would deal with "word learning". First, we do away with the notions of "lexical" and "word" acquisition: Under DM grounds, children are to acquire features, roots, and categorisers, as shown in section 2.4, as well as place these into the appropriate lists. List 1 lists features, bundles of features, roots, and categorisers, the complete inventory of any given language. List 2 stores associations between phonological exponents and the contents of syntactic terminal nodes. List 3 is where idiomatic meaning is stored as well as where general world knowledge interfaces with linguistic material.[8] Next, we present a hypothesis of what "word learning" looks like in DM.

3.1. The Hypothesis

The learnability problem delineated in section 2 – specifically, the observed decline in convergence of word-meaning associations within a computational model when exposed to BP utterances (Faria 2015) – underscores the necessity for revisions in cross-situational models of lexical acquisition. These revisions would not impact the core learning mechanisms advanced by each model, but rather entail the incorporation of a preliminary processing phase aimed at the extraction of morphemes from input utterances, a proposition substantiated by empirical investigations in language acquisition indicating children's abilities at discerning and producing morphemes from an early developmental stage (Resende 2021b). Furthermore, work on languages more morphologically complex than English and BP, such as polysynthetic languages, has shown that children's early productions are bare roots (and, therefore, morphemes) or attempts at extracting morphological material (Kelly *et al.* 2014; Allen 2017). The representational adaptations called for – that children represent linguistic input segmented into morphemes, not into words – align with the conceptual framework of non-lexicalist theories of grammar, such as DM, briefly outlined in the previous section. We propose a hypothetical guiding principle for lexical acquisition which allows children to learn morphemes and build words and sentences.

We contend that children by principle *seek to maximally segment the input data ("utterances") into its minimal components*. This principle sounds quite plausible, given that children will inevitably have to discover morphemes, regardless of whether a discrete lexical module is predicted by the chosen formal theory of grammar. Moreover, we contend that hints of this principle's existence are scattered around the literature. Let us now examine these clues in greater detail.

Albuquerque, Bezerra & Ferrari-Neto (2012) postulate that a theory of morphological acquisition has to include a description of input data such that the learner is able to determine and process linguistic units, as well as the attributes of both data and learner that ultimately lead to the successful acquisition of morphological material (strongly resonating with Atkinson's 1997: 91 description of the basic assumptions a learnability framework). The first aspect, a description of the data, is accounted for by the relationship between phonology, phrases, and word

8 Although we are not going to go as far as explaining how List 3 fits into lexical acquisition, we suggest that this is where the Third Factor (Chomsky 2005) and the third List interface.

segmentation, as discussed by Christophe *et al.* (2008). The second aspect, data and learner attributes, are found in Corrêa's (2009) hypothesis that formal features are correlated with the phonetic regularities found at phrasal borders.

Christophe *et al.* (2008) argue that lexical and syntactic learning are bootstrapped by the identification of prosodic phrases and of the regularities found within these phrases. Such prosodic phrases coincide with syntactic phrases, according to the authors, allowing the child to notice regularly occurring grammatical words, which in turn serve as a signal to segment and classify the accompanying content words. These prosodic phrases in essence define a search space in which, as the argument goes, more frequent phonetic material (function words, *e.g.* determiners and prepositions) allow for the extraction of the comparatively less frequent content words. Thus, the regular appearance of determiners such as *o* and *a* (respectively masculine and feminine 'the') enable the identification and classification of content words in *o menino, o gato, a garrafa, a lâmpada* ('the boy', 'the cat', 'the bottle', 'the light bulb'), *etc*.

Seeking to reconcile psycholinguistic processing theories with the acquisition of the formal features described in MP (Chomsky 1995), Corrêa (2009) posits that a correlation exists between sound regularities in Phonological Form and the formal features comprising the Numeration of a given derivation. Given the continuity between the syntactic derivation and the forms sent to the interfaces, the learner can infer the existence of a formal feature from phonological regularities. The author hypothesises that rhythmic patterns serve as indicators of whether the language adheres to a head-complement or complement-head structure, thus playing a role in establishing directional features. Additionally, a feature encoding the lexical/functional dichotomy would be assigned to initial lexical items, based on their distributional and phonological properties. These minimal elements are deemed sufficient to initialise the computational system, thereby marking the child's entry into syntax.

From the standpoint of a non-lexicalist hypothesis of morphological acquisition, these ideas may be operationalised in the following way. From the standpoint of the 'lexical' learner, the two most relevant properties of the input data are its phonological distribution and patterns. Within a given phonological phrase, the syllabic sequences comprising the input data are either frequent or infrequent, relative to one another. Christophe *et al.*'s (2008) model explains how a child, having heard *o menino, o gato, o cachorro*, ('the boy', 'the cat', 'the dog') is able to determine the pattern *o* __, where the gap represents a variable content word. Taking their ideas a step further, one could design a learner that listens for regularities (*cf.* Yang 2016: 43) at all borders of phonological phrases, learning to the identification of the pattern *o* __*o*, that is, of the determiner *o* and of the ending vowel *o*, a common morphological feature of languages such as BP. The gap represented by the dash allows the child to learn roots – the variable, "content" units of an utterance. Further iteration would lead to the discovery of more frequent material, as in the case of the set *o menininho, o gatinho, o cachorrinho*, leading to the template *o* __*inh-o*, where *-inh-* is a degree morpheme expressing diminutive in BP.

The most *frequent* material, besides being readily mappable to rudimentary functional features (Corrêa 2009), allow for the identification of the *infrequent* material, which is often mapped to roots – in the cases above, √MENIN, √GAT, √CACHOR. This observation would suggest a potential reinterpretation of the goals of classic lexical learning theories (Markman 1990; Bloom 2000; Golinkoff & Hirsh-Pasek 2000, among many others) as models primarily focused on explaining *root* *acquisition*. A morphological analysis of the data could potentially make computational models more cross-linguistic. Determiners like 'the' precede less frequent elements, which are typically roots; however, this pattern does not consistently apply to BP, where phrases like *o menino* and *a menina* ('the boy', 'the girl') and *o jarro* and *a jarra* ('the clay jug', 'the glass jug') end with a final vowel, effectively doubling the number of mappings the model has to establish. Nevertheless, if in a model like Faria (2015) the linguistic input data were segmented into atomic units, such as the root *menin-*, the ending vowels *-o* and *-a*, and the determiners *o* and *a*, the lexical sparsity in BP utterances would more closely resemble that of English, arguably increasing the performance of the model.

The learning mechanism proposed above is able to find regular and irregular sound material which respectively map into affixes and roots. With these primitives, the child can now structure their lexicon in terms of DM. As reviewed above in section 2.4, this theoretical model of grammar subdivides the lexicon into three lists. As seen, List 1, or the strict Lexicon, stores the primitives of syntactic derivation, consisting of features and bundles of features, categorisers, and roots. For example, a word like *gatinho* ('kitten', 'small cat'), could be represented[9] by the set {√GAT, $n_{[+\text{MASC}]}$, [degree: DIM]}. List 2, or Vocabulary, houses the associations between phonological exponents and the terminal nodes occupied by each of the elements above or combinations thereof. In the example above, the corresponding Vocabulary entries could be /gat/ ↔ √GAT, /iŋ/ ↔ [degree: DIM], /o/ ↔ $n_{[+\text{MASC}]}$. The task of lexical acquisition the child faces is, thus, to instantiate a feature (and occasionally a categoriser) in List 1 whenever sound regularities are found within prosodic units (Corrêa 2009) and to record the sound-feature association in List 2. By the same token, the accompanying (infrequent) sounds are going to trigger the postulation of a new root in List 1, as well as the corresponding Vocabulary entry in List 2. These are not finalised linguistic elements, being subject to scrutiny by the cross-situational procedures previously described (Siskind 1996; Fazly, Alishahi & Stevenson 2010; Trueswell et *al.* 2013; among many others).

Having a description of what lexical learning entails in a non-lexical theory of language, we can turn to the goal of evaluating DM against some facts of acquisition. Namely, we present under DM a formal treatment of the U-shaped curve (section 3.2), as well as an account for linguistic change (section 3.3).

3.2. Distributed Morphology as a well-suited model for explaining the U curve in Brazilian Portuguese

The fact that children's early irregular verbs go through a phase of overregularisation (*e.g. trazi* 'bringed' instead of *trouxe* 'brought') after having produced the adult-like forms has been described for languages such as English (Pinker & Prince 1988) and BP (Maldonade 2003; Lorandi 2010; Figueira 2010). Also known as the U-shaped curve of irregular verb acquisition, this phenomenon serves as a valuable benchmark for evaluating the predictive capacity of linguistic models regarding the developmental stages of children's grammar. We pull an example from Araújo-Adriano & Beraldo

9 The essential point here is not to claim these descriptions as definitive representations of *gatinho* under DM, since finding the feature inventory and syntactic structure of a language like Brazilian Portuguese remains an ongoing field of research. Rather, the chosen set of features, root, and categoriser merely exemplifies one potential analysis among many.

(2023) to illustrate how linguistic theories can be ruled out by the yardstick of acquisition.

Three approaches are critically reviewed and found insufficient in explaining the U curve (Araújo-Adriano & Beraldo 2023). The first approach (Lorandi 2010) explains overregularisation based on Optimality Theory by positing that children prioritise root fidelity, producing forms *trazi* ('bringed'), whereas adults prioritise non-fidelity, producing *trouxe* ('brought'). However, this approach fails to account for why children eventually abbandon root fidelity in favour of target forms. The second approach (Wuerges 2014) examined the role of input frequency in verb acquisition. Using Yang's (2002) Variational Model, the author successfully predicted which irregular verbs are more prone to overregularisation based on input frequency. Although the Variational Model addresses the acquisition of verbs in the presence of input noise, it is unable to explain the acquisition of exceptions – a limitation later acknowledged by Yang (2016). Given that these approaches do not adopt the framework of Distributed Morphology, they will not be further examined, as they do not address this framework's explanatory adequacy regarding language acquisition.

The third approach (Takahira 2013) does adopt Distributed Morphology and argues that children's overregularisated productions lack movement from *v* to T, such that the terminal nodes root √FAZ and Tense/Agree containing the [PRF.PAST, 1, SG] features are available for the insertion of the morphemes /faz/ and /i/, ultimately resulting in the production of /fazi/. Takahira contends that this absence of movement can be attributed to children's incomplete acquisition of the Minimise Exponence principle (Siddiqi 2009) which, in adult grammar, triggers the movement so that the most minimal exponent /faz/ can be inserted. However, this explanation falls short in elucidating why children's earlier productions, that is, before overregularisation, often appear to conform to the Minimise Exponence principle, yielding forms such as /fiz/. Moreover, conflicting evidence (Santos & Lopes 2017) challenges the assertion that children's grammar would not move *v* to inflection.

Hence, the question arises regarding the compatibility of non-lexicalist models, such as Distributed Morphology, with the U-shaped curve phenomenon. Traditionally, the onset of overregularisations in the past Tense was seen as evidence of the acquisition of the past-Tense forming rule. Overregularisations were simply an instance of failure in accessing the memorised irregular form, which normally blocks the regular rule from being applied to irregular verbs (Pinker & Prince 1988; Marcus *et al.* 1992). On the other hand, Yang (2016: 29) presents an alternative perspective, proposing that all verbs are produced by morphological rules. Thus, irregular verbs such as *brought* and *sought* are produced by the assignment of *bring* and *seek* to the "ought" rule, whereas all regular verbs are assigned to the "add -ed" rule. When producing a past form, the target verb is checked against a list of exceptions; if it does not match any exceptional entries, then the "elsewhere rule" (Berwick 1985) is applied. This means that the de facto regular rule is always the last resort, consistent with the observation that irregular forms are retrieved faster than their regular counterparts (p. 49).

Yang's (2016) notable contribution lies in his mathematical analysis of rule productivity. According to Yang (2016: 48), regular rules are inherently limited in their ability to accommodate exceptions, lest the regular rule become unproductive itself. This constraint is quantified by the Tolerance Principle (TP), which dictates that the number of exceptions must remain below a certain threshold.[10]

Interestingly, as a result of TP, the U curve becomes a matter of productivity. Initially, irregular verbs, being more prevalent in (child-directed) language data, are posited as the primary rules, while regular verbs, being comparatively less frequent, are treated as exceptions. As a child's vocabulary expands, this balance is tipped over, and irregular verbs transition to the status of exceptions to the regular rule. This transition creates a window of opportunity for irregular verbs to be temporarily reanalysed, being subject to the application of the regular rule, before being accurately listed as exceptions and produced correctly.

A model like Distributed Morphology, however, operates without a specialised lexicon and lacks the capacity to represent rules as expected by the Tolerance Principle. At first glance, this might suggest that these approaches are incompatible. Nevertheless, we propose that some of the same learning mechanisms observed in Yang (2016), specifically inductive pattern finding and the TP, are at play within a non-lexicalist framework. These mechanisms offer insights into the three stages of language acquisition: The use of adult-like forms (Stage 1), the phase of overregularisation (Stage 2), and the eventual attainment of the target grammar (Stage 3). We illustrate how the notion of the Vocabulary Item (see section 2.1 above) can replace rule-based morphological acquisition by examining the development of the past Tense in BP, with implications that naturally extend to other instances of irregularity.

In Stage 1 of past-Tense acquisition, children are faced with the prevalence of irregular verbs. Since VI must be memorised by speakers (Halle 1997: 128), during this phase, children engage in the memorisation of morphemes (in DM, the phonological side of VIs) and the combinations of verb roots with associated features, as illustrated in example (6a). Stage 2 is characterised by the emergence of the regular rule, which becomes evident as the number of exceptions decreases with the steady acquisition of a greater number of regular verbs. This transition prompts children to reanalyse their previously acquired Vocabulary Items (VIs) into more concise forms, associating the [PRT.PST, 1, SG] with a single morpheme /i/, as shown in example (6b). It is noteworthy that Minimise Exponence continues to influence children's derivations in both stages, depending on their inventory of VIs. The progression to Stage 3, wherein verbs are produced in accordance with the target adult grammar (6c), is achieved as irregular forms gradually supersede the application of regular VIs. A summary of this process, as well as proposed derivations, can be seen in Table 1.

(6) a. Stage 1: Memorisation of adult-like forms

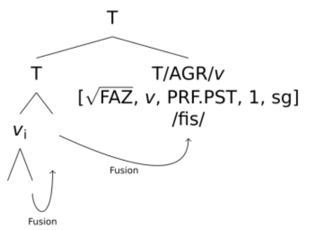
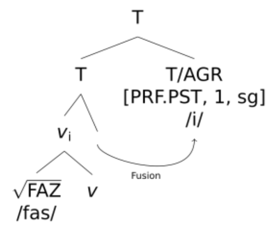
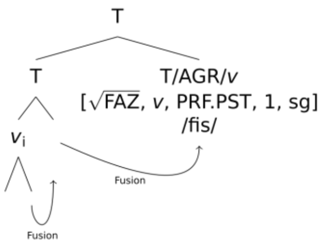
/'fis/	↔	[√FAZ, *v*, PRF.PST, 1, SG]	(*fiz*, 'I did')
/ko.'mi/	↔	[√COM, *v*, PRF.PST, 1, SG]	(*comi*, 'I ate')
/dor.'mi/	↔	[√DORM, *v*, PRF.PST, 1, SG]	(*dormi*, 'I slept')

[10] More specifically, the number of exceptions *e* must be smaller than or equal to the number of items governed by the regular rule *N* divided by its natural logarithm ln *N* (Yang 2016: 64).

b. Stage 2: Overgeneralization
/fas/ ↔ [√FAZ]
/kom/ ↔ [√COM]
/dorm/ ↔ [√DORM]
/i/ ↔ [PRF.PST, 1, SG]

c. Stage 3: Target language
/'fis/ ↔ [√FAZ, *v*, PRF.PST, 1, SG]
/'kom/ ↔ [√COM]
/dor'm/ ↔ [√DORM]
…
/i/ ↔ [PRF.PST, 1, SG]

Table 1. Stages in the Acquisition of Irregular Verbs

Stage 1	**Stage 2**	**Stage 3**
Vocabulary Items contain roots and features	*Roots and affixes are listed as separate Vis*	*Irregular verbs revert to being a singleton VI*
		

Source: Own work

This analysis seems to be consistent with the literature suggesting that morphemes, rather than words, are at the basis of lexical acquisition.[11] It also supports the present proposal that children look for the smallest units of language, which leads to the eventual reanalysis of the Vocabulary item set from Stage 1 to Stage 3. Finally, Araújo-Adriano & Beraldo (2023), the case study outlined above, shows how a non-lexicalist theory of grammar – Distributed Morphology, in this case – is well equipped with the descriptive means necessary to explain the processes involved in lexical acquisition, both in terms of the structures generated at all stages of the U-curve as well as the evolving lexicon.

3.3. Distributed Morphology as a well-suited model for explaining diachronic change in Brazilian Portuguese

Lightfoot's (1979) groundbreaking work has established a crucial link between diachronic change and language acquisition. This paradigm-shifting model for language change challenged the previous proposals, such as the one envisaged by

[11] The actual VIs are likely more complex than those in (6), given that BP verbs are associated with classes that must be acquired between Stage 2 (6b) and Stage 3 (6c). We have deliberately omitted these classes from our analysis, as their status remains an open question within Distributed Morphology (see footnote 6). Furthermore, including them would imply a theoretical stance on their acquisition, which falls beyond the scope of this paper.

Klima (1964), in which diachronic change was taken to be generated by different grammatical system rules. Under the Transformational Generative Grammar approach (Chomsky 1957), sentences were generated by the interplay between Phrase Structure rules and Transformational rules, some of which had a specific order. For instance, a rule named Case Marking had to be applied in a given structure before Wh Movement so that the dislocated WH-constituent is already marked with an Accusative Case in a local domain. Conversely, the wh element would not be able to be marked with Accusative Case if the Wh Movement rule was applied first. Within this background, Klima's proposal for diachronic change was that change would occur as a result of different rule orders across time. For instance, a grammar of period 1 (G1) whose grammatical rules are ordered as *rule a* > *rule b* > *rule c* generates a different system from another grammar of period 2 (G2) that orders the same rules as *rule a* > *rule c* > *rule b*. This proposal, however, allowed for any possible change to happen, an undesired result, as pointed out by Lightfoot himself:

> If the constraints on grammars are so loose that G1 or G2 can take on an unlimited number of forms, then G1 may differ from G2 in an unlimited number of ways and we have no method of distinguishing possible from impossible changes, which I take to be a central task of any theory of linguistic change (Lightfoot 1979: 14).

Considering this flaw, Lightfoot (1979) proposes that language acquisition triggers language change. This link comes from the idea that children will always acquire a language based on their parents', regardless of the historical period. Still, this acquisitional process is far from perfect, for children do not have direct access to their parent's grammar: If acquisition were always perfect, in a one-to-one relation to children's parents, "changes within a population would seemingly never occur" (Niyogi & Berwick 1995: 1). Based on Andersen's (1973) work on the acquisition of Russian phonology, Lightfoot then proposes that a system changes when a generation of speakers entertains a slightly different grammar from the previous generation's output, where "small differences in output may result from large differences in the grammar, and vice-versa" (Lightfoot 1979: 147). However, when the output is the same, a change can also occur because children might have abducted[12] a different system that generates the exact input string. When this happens, a change occurs: Children's 'mistakes' today are tomorrow's linguistic changes (Andersen 1973; Lightfoot 1979).

[12] Here, we take abduction to be an explanatory hypothesis as a conclusion. Given some evidence, one can entertain some hypotheses, but they do not necessarily converge with reality. Andersen (2017) explains this idea *via* the following narrative.

> We were driving through Centerville and decided to drop in on our friends on Elm Street. When we got there, to our surprise, there was no light on in their house [C]. Our first thought was that they had gone to bed [A1]. But it was barely past 9 p.m. They might be out on the town [A2] or visiting friends [A3]. Or perhaps they were away on vacation [A4] (Andersen 2017: 304)

From the narrative, A1, A2, A3, and A4 are all possible conclusions to C.

The abductive rationale applied to language changes was dubbed the Z-model of linguistic change (*cf.* (7) below), where Generation 1 offers Generation 2 some input – *Corpus* 1 – based on their Grammar 1, but Generation 2 might abductively analyse *Corpus* 1 and attribute to it a different structure. Through this abductive analysis, Grammar 2 differs slightly from Grammar 1, generating a *Corpus* 2 that differs from the previous one. In this sense, "language acquirers converge on a grammatical system which differs in at least one […] value from the system internalised by the speakers whose linguistic behaviour provides the input to those acquirers (Roberts 2007: 227)." A linguistic example of this abductive analysis comes from the synthetic Romance future from the analytic Latin future.

In Classic Latin, the synthetic future Tense was formed by the verbal root and suffixes, such as -BO: EDUCABO 'educate.FUT.' In Vulgar Latin, an alternative to *educabo* was a periphrastic future Tense, formed by infinitive + HABERE ('have') inflected to the present Tense: EDUCARE HABEO 'educate.INF have.PRS.' In Archaic Portuguese and other Romance Languages, the periphrasis with the infinitive + HABEO, which surpassed the synthetic form in -BO, underwent phonological processes, *viz.*, *habeo* > *haeo* > *hae* > *hai* > *hei* (Câmara Jr. 1976; Nunes 2003), giving rise to the the following periphrasis infinitive + *hei* (PT. *educar hei* 'educate.INF have.PRS = I will educate') or *hei* + infinitive (PT. *hei educar* 'have.PRS educate.INF = I will educate'). Since the order of the auxiliary and the infinitive form was free (Nunes 1992), some writers stylistically opted for *hei comer* 'I will eat' or *hei de comer*, with the intervening preposition, or even *comer hei* 'I will eat.' Exposed to sentences with infinitive followed by HABERE, *e.g. comer hei*, *i.e.*, an analytic structure, children could abductively analyse the structure to mark futurity as a singleton, that is, *comer* + suffix (h)ei, as a synthetic structure, *viz.*, *comerei* ('I will eat'):

(7) Generation (G1) → *Corpus* 1 (*comer hei* [ko.'mer.ej])
↙
Generation (G2) → *Corpus* 2 (*comerei* [ko.me.'rej])

As seen, Lightfoot implicitly highlights the crucial role of language acquisition in driving linguistic change through the lens of Generative Grammar. His proposal that change is due to erroneous analysis by children underscored the symbiotic relationship between language acquisition and diachronic syntax, illuminating how the dynamic interplay between these phenomena engenders the rich tapestry of different linguistic synchronies observed across different historical periods.

Considering how the mechanisms of language acquisition shape and are shaped by the evolving linguistic landscape, any linguistic theory for language acquisition must also account for language change. Hence, if DM is on the right track, the expectations are that it would successfully deal with language acquisition, In the previous sections we showed it does; by extension, DM would also account for language change. That is, under the assumption that 'word' does not have any special status regarding language acquisition, the notion of 'word' would not receive any special treatment concerning diachrony either, promoting DM as a well-suited model for explaining diachronic change. In this section, we entertain this idea: We show changes smaller than words, affecting affixes (Bacheschi 2006; Olsen 2014; Resende & Ilari 2020) and even smaller entities, such as features, in the case of the well-studied reduction in the BP inventory of verbal present morphemes (Duarte 1993).

Observing changes in domains other than sentences, one finds changes in units smaller than words. The empirical ground that supports this claim comes from changes that are not found in a word as a whole but in smaller units. In BP, the suffix *-ite*, stemming from the Greek suffix -ITIS, is a "suffix element of composition indicating disease or inflammation of the organ or anatomical structure indicated in the root",[13] that is, a suffix traditionally denoting "inflammation of x," where "x" refers to an organ. For instance, the Portuguese word *hepatite* ('hepatitis'), from Greek HÊPAR ('liver') + -ITIS ('inflammation'), means inflammation of the liver, *meningite* ('meningitis'), inflammation of the meninges, among others. In a more recent development,[14] however, this suffix has taken on a broader meaning beyond bodily organs associated with certain afflictions, vices, an "excessive tendency to x," "obsession with," such as *paixão* + -ite (PT. *paixonite* 'passion-itis = excessive passion towards someone'), *diploma* + -ite (PT. *diplomite* 'certificate-itis = obsession with certificates'), *frescura* + -ite (PT. *frescurite* 'fuss-itis = excessive fuss'), *preguiça* + -ite (PT. *preguicite* 'laziness-itis = excessive laziness') (*cf.* also Bacheschi 2006: 56; Resende & Ilari 2020: 257)[15].

Another example of change in a domain smaller than words comes from the pen of Sandmann (2010: 77). The author shows that the suffix *-ice* is the "prototipic depreciative morpheme" in BP, as in *mesquinho* + -ice (PT. *mesquitice* 'stingy-ice = excessive reluctance to spend money'), *gramática* + -ice (PT. *gramatiquice* 'grammar-ice = excessive or pedantic focus on grammar rules'), *moderno* + -ice (PT. *modernice* 'modern-ice = uncritical embrace of modern trends and fashions'), *brega* + -ice (PT. *breguice* 'tacky-ice = the use of clothes, objects, and behaviours considered of poor taste'), associated only with depreciative roots, given that roots that express positivity cannot be merged with this suffix: **lindo* + -ice (PT. *lindice* 'beautiful-ice'), **alegre* + -ice (PT. *alegrice* 'happy-ice'). However, according to Sandmann (2010), in the past, -ice did not have a depreciative reading, meaning "the way of being x," where "x" refers to a quality, like *menino* + -ice (PT. *meninice* 'boy-ice = childishness'), *meigo* + -ice (PT. *meiguice* 'sweet-ice = quality of being sweet, gentle, or affectionate'), *mineiro* + -ice (PT. *mineirice* 'miner-ice = the way of acting of someone who lives in Minas Gerais, a Brazilian state'), the latter added to the Aurélio dictionary as a synonym to *mineirismo* and *mineiridade* in mid 1988 (*cf.* Sandmann 1988: 50). Considering that it was the suffixes *-ite* and *-ice*, irrespective of their root, that assumed a new meaning, we can entertain the hypothesis that the mechanisms of change reveal that the notion of word lacks any privileged status within the diachronic realm as well (*cf.* also Resende & Ilari 2020: 257–258). Drawing on this discussion, we apply this hypothesis to other instances of diachronic change beyond individual words and consider this proposal in relation to smaller linguistic units, such as features.

The literature on BP has widely acknowledged that the verbal paradigm underwent a reduction. Investigating the emergence of null subjects, Duarte (1993) suggests that the inflection paradigm had six forms, but it was oversimplified to four or, according to Costa & Figueiredo Silva (2006: 99), even two forms in some BP

---

13 Porto Editora – *-ite* in *Dicionário Infopédia de Termos Médicos*. Porto: Porto Editora. Available on: <https://www.infopedia.pt/dicionarios/termos-medicos/-ite>. Access on May 10, 2024.

14 In a rapid survey, we searched on Google for the terms 'meningite' and 'paixonite' in a span of time between 1960 and 1990. We found 258 occurrences for the former against 4 for the latter.

15 This affix change from 'disease,' 'inflammation' to 'addiction of,' 'abnormal excess of' seems to have also occurred in English: *computeritis*, *cellphoneitis*, *facebookitis* (*cf.* Olsen 2014: 36).

dialects. In Table 1, one observes that in the past, *circa* 1845 in Duarte's data, BP had six verbal endings: /u/ as in *cant/u/* ('sing.1.SG.PRS'), /s/ as in *canta/s/* ('sing.2.SG.PRS'), /ø/ as in *canta/ø/* ('sing.3.SG.PRS'), /mus/ as in *canta/mus/* ('sing.1.PL.PRS'), /is/ as in *canta/is/* ('sing.2.PL.PRS'), and /N/ as in *canta/N/* ('sing.3.PL.PRS'). After various independent changes, such as the grammaticalisation of *a gente* (lit. 'the people') and *você* ('you') from both pronominal forms of 3rd person to 1st plural and 2nd singular pronominal forms, respectively, the neutralisation between *tu* and *você*, in favour of the latter, and other neutralisation between *a gente* and *nós*, where *a gente*, after 1970, gains ground, the paradigm underwent simplification, there being only two endings that differentiate person in the verb template: /u/ as in *cant/u/* ('sing.1.SG.PRS') *versus* /ø/ as in *canta/ø/* ('sing.2/3.SG.PRS' or 'sing.1/2/3.PL.PRS'). In this case, the verb does not show any agreement mark of number other than the 1st person singular.

Table 2. BP verbal paradim change from six forms to two forms

	A six-form paradigm	**A four-form paradigm**	**A two-form paradigm**[16]
eu	cant-o /kan.tu/	cant-o /kan.tu/	cant-o /kan.tu/
tu *você*	canta-s /kan.tas/ -	canta-(s) /kan.ta(s)/ canta-ø /kan.ta/	canta-ø /kan.ta/
ele, ela	canta-ø /kan.ta/	canta-ø /kan.ta/	canta-ø /kan.ta/
nós *a gente*	canta-mos /kan.ta.mus/ -	canta-mos /kan.ta.mus/ canta-ø /kan.ta/	canta-ø /kan.ta/
vós *vocês*	canta-is /kan.tais/ canta-m /kan.taN/	- canta(m) /kan.ta(N)/	canta-ø /kan.ta/
eles, elas	canta-m /kan.taN/	canta-m /kan.ta(N)/	canta-ø /kan.ta/

Source: Adapted from Duarte 1993; Costa & Figueiredo Silva 2006

Assuming that language acquisition and, consequently, diachronic change do not prioritise the notion of the word, it can be argued that the changes in the verbal paradigm were a result of units smaller than words. We argue that this change concerns feature underspecifications that phonologically realised the φ-features of 1st, 2nd, and 3rd persons singular/plural.

Under the Distributed Morphology approach, agreement markings have no syntactic role. Therefore, it is assumed that all φ-features of the subject are copied onto an Agr(eement) node that is inserted post-syntactically, due to a well-formedness

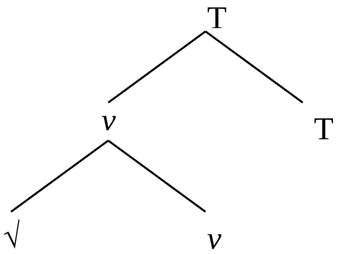
---

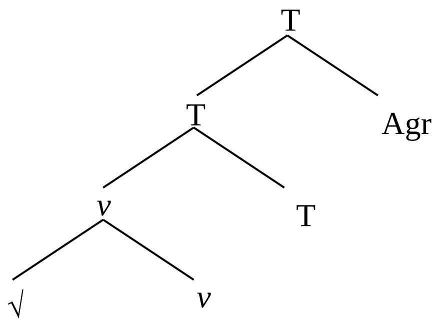
16 An attendee of the event "30 Years Distributing Morphology from North to South (DM30)" reminded us that in the imperfect past Tense, the paradigm is even more reduced, in a one-form paradigm:

(i) a. Eu comia ('I ate')
b. Tu/você comia ('you.SG ate')
c. Ele comia ('he/she ate')
d. Nós/a gente comia ('we ate')
e. Vocês comia ('you.PL ate')
f. Eles comia ('they ate')

condition in Morphological Structure, not present during syntax (Embick & Noyer 2007), as the following representations depict. Some languages display an additional operation that fuses T and Agr nodes, which is the case of Portuguese in some contexts (*cf.* Bassani & Lunguinho 2011).

(8) a. Syntax b. Morphological Structure





In this paper, we adopt Halle's (1997) proposal for the information of person in terms of binary features, under which person is a combination of ±AUTHOR and ±PARTICIPANT. In addition to that, in the Agr node there must also be encoded information about number, ±PLURAL. Thus, BP φ-features in T are a result of the combination in (9).[17]

(9) φ-features in Agr
a. [+AUTHOR,+PARTICIPANT, ±PLURAL]: 1st person
b. [-AUTHOR,+PARTICIPANT, ±PLURAL]: 2nd person
c. [-AUTHOR, -PARTICIPANT, ±PLURAL]: 3nd person

After Agr node insertion (*cf.* (8b)) and additional operations have taken place, for instance, Fusion, phonological exponents are provided in List 2. In the present Tense (PT. *canto* 'sing.1.SG.PRS'), Fusion occurs between *v*, Agr, and Tense (*cf.* also Bassani & Lunguinho 2011). Sequentially, the following phonetic material is inserted in one single node, in the case of 1, singular, present: /u/ ↔ [+AUTHOR,+PARTICIPANT, -PLURAL]/[+PRESENT] (*cf. cant/u/* 'sing.1.SG.PRS').

Concerning the verbal paradigm change in BP discussed previously, under the initial paradigm, in List 2, BP speakers had a one-to-one phonological correspondence to express φ-features in the verb, that is, there were six phonological exponents dedicated to the realisation of six φ-feature combinations in Vocabulary[18].

(10) Six-form paradigm

/u/	↔	[+AUTHOR, +PARTICIPANT, -PLURAL]
/s/	↔	[-AUTHOR, +PARTICIPANT, -PLURAL]
/ø/	↔	[-AUTHOR, -PARTICIPANT, -PLURAL]
/mus/	↔	[+AUTHOR, +PARTICIPANT, +PLURAL]
/is/	↔	[-AUTHOR, +PARTICIPANT, +PLURAL]
/N/	↔	[-AUTHOR, -PARTICIPANT, +PLURAL]

---

17 In Halle's (1997) proposal, a fourth combination is also possible, *viz.* [+AUTHOR, -PARTICIPANT, ±PLURAL]. That would encompass Walbiri, a language that encodes a referent as "I and someone else, but not you."

18 For an alternative proposal regarding the phonological exponents that realise φ-features in BP, *cf.* Nunes (2020) and Kato, Martins & Nunes (2023).

The inclusion of *a gente* ('the people,' meaning 'we') and *você(s)* ('you[.PL]') in the pronominal paradigm altered children's evidence of the pairing of features and phonetic realisations: The two third-person pronouns refer to the first- and second-person, associated with [+AUTHOR, +PARTICIPANT, ±PLURAL] and [-AUTHOR, +PARTICIPANT, ±PLURAL], respectively, but the phonetic evidence available to children, /ø/, is the one associated with the third person [-AUTHOR, -PARTICIPANT, -PLURAL]. As a consequence, the new φ-feature system started exhibiting certain underspecified features, namely [AUTHOR] and [PARTICIPANT], given children's different evidence of the pairing of features and phonological realisations.

In the novel system, attested by Costa & Figueiredo Silva (2006), there is no one-to-one phonological correspondence to express φ-features, as in (10). Instead, the same phonological matrix, *viz.* /ø/, can realise features that are underspecified for AUTHOR and PARTICIPANT.

(11) Two-form paradigm
a. /u/ ↔ [+AUTHOR, +PARTICIPANT, -PLURAL]
b. /ø/ ↔ [±PLURAL]

This is evident in all persons and numbers except for 1st person singular, which remains the most specified (but see footnote 14): Features [+AUTHOR, +PARTICIPANT, -PLURAL] are still realised by /u/. For other persons and numbers, however, the same phonological content is realised regardless of the [AUTHOR] and [PARTICIPANT] features: [±PLURAL] ↔ /ø/ as in *canta/ø/*. This amounts to saying that if in syntax there is, for instance, a DP *Os menino*, as in (12), associated with [-AUTHOR, -PARTICIPANT, +PLURAL], the terminal node in Agr, after DP's φ-feature is copied onto Agr node, cannot receive the Vocabulary insertion in (11a), given that -PLURAL does not appear at the terminal node. However, given Underspecification (Halle & Marantz 1994: 276), the Vocabulary Item in (11b) [±PLURAL] is a subset of the bundle of features [-AUTHOR, -PARTICIPANT, +PLURAL] provided by the terminal node in syntax.

(12) (adapted from Costa & Figueiredo Silva 2006: 100)

[Os	menino$_{[+\text{AUTHOR}, +\text{PARTICIPANT}, +\text{PLURAL}]}$]	come-ø	o	doce
The.PL	boy	eat.PRS-3SG	the	candy

If this proposal is accurate, a morphological model that does not prioritise 'word' as a fundamental unit also provides an explanation for how a linguistic system undergoes transitions from one state to another, in the present case, from a six-form to a two-form paradigm. As Lightfoot (1979: 47) pointed out, "small changes in E-language sometimes trigger new I-languages, with more farreaching consequences". In this case, we propose that the presence in E-language of 'a gente' and 'você', progressively more associated with [+AUTHOR, +PARTICIPANT, ±PLURAL] and [-AUTHOR, +PARTICIPANT, ±PLURAL], respectively, but paired with ø – the same exponent as the third person – triggered reanalysis of the φ-feature system in I-language. This reanalysis occurred as a consequence of there being no distinction between the pairing of features and phonological realisations. As a corollary, children rearranged their pairing, giving rise to a different system, namely, a two-form paradigm. Hence, considering that diachronic change and language acquisition can be analysed in complementary ways, and based on the proposal suggested, we presented empirical studies assuming Distributed Morphology, according to which the word –

and not just the sentence – is a product of syntactic derivation and its consequences for language acquisition.

4. Conclusion

In the beginning of section 2, we likened the relationship between acquisitional data and linguistic theory to a feedback loop. If, on the one hand, linguistic theory provides a representation for language and acquisitional input data, on the other hand, factors of language acquisition help evaluate a theory's accuracy. We revisited a computational model of lexical acquisition (Faria 2015) that implemented an algorithm to extract a lexicon from *corpora* of child-directed speech. A significant drop in performance occurred when the chosen *corpus* was changed from English data to Brazilian Portuguese data. This decline was attributed to the more complex morphology of Brazilian Portuguese words.

Drawing on research on the acquisition of polysynthetic languages (Kelly *et al.* 2014; Allen 2017) and empirical arguments against the special status of the word (Marantz 1997; Resende 2021a), we concluded that morphological words cannot be the unit of lexical acquisition. We then explored Distributed Morphology (DM; Halle & Marantz 1993), a non-lexicalist model of grammar, which we argued is better suited for explaining how words are learned. This is because DM describes the lexicon as a distributed system where roots, features, and categorisers are listed separately from the instructions mapping these elements into their phonological realisation.

Based on previous research on lexical and syntactic bootstrapping (Christophe *et al.* 2008; Corrêa 2009), we proposed that children segment input data into the smallest possible units, traditionally called 'morphemes.' These atomic units populate the lexicon as prescribed by DM and are subject to constant reanalysis during the acquisition process. By examining empirical evidence on lexical acquisition and a theory of grammar and acquisition, we completed the feedback loop suggesting that children follow a principle of maximal segmentation which, combined with the architecture of the lexicon from DM, explains not only the U-shaped curve in the acquisition of Brazilian Portuguese irregular verbs (Pinker & Prince 1988; Maldonade 2023; Araújo-Adriano & Beraldo 2023) but also the observed linguistic change (Lightfoot 1979) in the verbal system of that language.

Acknowledgments

We sincerely thank the Isogloss reviewers and editors for their insightful comments and suggestions, which helped to markedly improve the original manuscript.

References

Aikhenvald, Alexandra Y. 2012. Building words. In _____, *The languages of the Amazon*. Oxford University Press.

Albuquerque, Hertha, Bezerra, Gitanna, & José Ferrari-Neto. 2012. Percepção infantil da morfologia derivacional: Um estudo experimental sobre segmentação de morfemas em português brasileiro. *Signo y Seña* 22, 119–138.

Alcântara, Cíntia da Costa. 2003. *As classes formais do português e sua constituição: um estudo à luz da teoria da morfologia distribuída*. Ph.D. thesis, Porto Alegre: Pontíficia Universidade Católica do Rio Grande do Sul.

Alcântara, Cíntia da Costa. 2010. As classes formais do português brasileiro. *Letras de Hoje* 45(1): 5–15.

Allen, Shanley. 2017. Polysynthesis in the acquisition of Inuit languages. In M. Fortescue, M. Mithun, & N. Evans (eds.), *The Oxford handbook of polysynthesis*, 449–472. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199683208.013.25

Andersen, Henning. 2017. Abduction. In A. Ledgeway and I. G. Roberts (eds.), *The Cambridge handbook of historical syntax*, 301–321. Cambridge Handbooks in Language and Linguistics. Cambridge: Cambridge University Press.

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49(4): 765–793.

Araújo-Adriano, Paulo Ângelo, & Rafael Luis Beraldo. 2023. Implementing the U-shaped curve in Distributed Morphology and as a by-product of the Third Factor. *Journal of Portuguese Linguistics* 22(1). https://doi.org/10.16995/jpl.8961

Atkinson, Martin. 1997. Learnability. In P. Fletcher & M. Garman (eds.), *Language acquisition: Studies in first language development*, 2nd ed., 90–108. Cambridge: Cambridge University Press.

Bacheschi, Celso. 2006. *Os valores expressivos dos afixos na norma urbana culta de São Paulo.* Master's dissertation, São Paulo: Pontifícia Universidade Católica.

Bagetti, Tatiana, & Letícia Sicuro Corrêa. 2010. The early recognition of verb affixes: Evidence from Portuguese. *On-Line Proceedings of the 35th Annual Boston University Conference on Language Development*. 35th Annual Boston University Conference on Language Development, Boston.

Bassani, Indaiá de Santana, & Marcus Vinicius Lunguinho. 2011. Revisitando a flexão verbal do português à luz da Morfologia Distribuída: Um estudo do presente, pretérito imperfeito e pretérito perfeito do indicativo. *Revista Virtual de Estudos da Linguagem* 5(9): 199–227.

Beraldo, Rafael Luis. 2020. *Computational models of lexical acquisition: Surveying the state of the art.* Master's dissertation, Campinas: Universidade Estadual de Campinas. https://repositorio.unicamp.br/acervo/detalhe/1129285

Berwick, Robert C. 1985. *The acquisition of syntactic knowledge*. The MIT Press Series in Artificial Intelligence. Cambridge, MA: MIT Press.

Bloom, Paul. 2000. *How children learn the meanings of words*. Cambridge: The MIT Press.

Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge, MA and London: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Câmara Júnior, Joaquim Mattoso. 1976. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Chomsky, Noam. 1988. *Language and problems of knowledge: The Managua lectures*. Cambridge, MA: The MIT Press.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Current Studies in Linguistics 28. Cambridge, MA: The MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1): 1–22. https://doi.org/10.1162/0024389052993655

Christophe, Anne, Séverine Millotte, Savita Bernal, & Jeffrey Lidz. 2008. Bootstrapping lexical and syntactic acquisition. *Language and Speech* 51(1–2): 61–75. https://doi.org/10.1177/00238309080510010501

Coelho, Claudia Souza, & Paulo Ângelo Araújo-Adriano. 2021. Em favor de raízes sem diacríticos. *Caderno de Squibs: Temas em Estudos Formais da Linguagem* 7(2): 44–52. https://www.periodicos.unb.br/index.php/cs/article/view/40408

Corrêa, Letícia Sicuro. 2009. Bootstrapping language acquisition from a minimalist standpoint: On the identification of φ-features in Brazilian Portuguese. In A. Pires, & J. Rothman (eds.), *Minimalist inquiries into child and adult language acquisition*, 35–62. Studies on Language Acquisition. Berlin; New York: Mouton de Gruyter. https://doi.org/10.1515/9783110215359.1.35

Costa, João, & Maria Cristina Figueiredo-Silva. 2006. Notas sobre a concordância verbal e nominal em português. *Estudos Linguísticos* 35: 95–109.

Duarte, Maria Eugenia Lammoglia. 1993. Do pronome nulo ao pronome pleno. In I. Roberts, & M. Kato (eds.), *Português brasileiro: Uma viagem diacrônica*, 107–28. Campinas: Editora da Unicamp.

Embick, David, & Rolf Noyer. 2007. Distributed Morphology and the syntax-morphology interface. In G. Ramchand, & C. Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 289–324. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199247455.013.0010

Faria, Pablo. 2015. A computational study of cross-situational lexical learning of Brazilian Portuguese. In *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning*, 45–54. Lisbon, Portugal: Association for Computational Linguistics. https://doi.org/10.18653/v1/W15-2408

Fazly, Afsaneh, Afra Alishahi, & Suzanne Stevenson. 2010. A probabilistic computational model of cross-situational word learning. *Cognitive Science* 34(6): 1017–1063. https://doi.org/10.1111/j.1551-6709.2010.01104.x

Ferrari-Neto, José, & Marcus André Ferraz de Lima. 2015. Aquisição da morfologia flexional verbal em português brasileiro – um estudo com dados de compreensão. *Revista Prolíngua*, 10(1): 106–120.

Figueira, Rosa Attié. 2010. O que a investigação sobre o erro na fala da criança deve a Saussure. *Cadernos de Estudos Lingüísticos* 52(1): 115–144. https://doi.org/10.20396/cel.v52i1.8637206

Fisher, Cynthia, Hall, D. Geoffrey, Rakowitz, Susan, & Lila Gleitman. 1994. When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, 92: 333–375. https://doi.org/10.1016/0024-3841(94)90346-8

Guimarães, Maximiliano. 2017. *Os fundamentos da teoria linguística de Chomsky*. Petrópolis: Vozes.

Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition*, 1(1): 3–55. https://doi.org/10.1207/s15327817la0101_2

Gold, E. Mark. 1967. Language identification in the limit. *Information and Control*, 10(5): 447–474. https://doi.org/10.1016/S0019-9958(67)91165-5

Golinkoff, Roberta, & Kathy Hirsh-Pasek. 2000. Word learning: icon, index or symbol? In A. L. Woodward, R. M. Golinkoff, L. Bloom, K. Hirsh-Pasek, L. B. Smith, N. Akhtar, M. Tomasello, & G. Hollich (eds.), *Becoming a word learner: A debate on lexical acquisition*, Counterpoints: Cognition, Memory, and Language. Oxford: Oxford University Press.

Halle, Morris, & Alec Marantz. (1993). Distribuited Morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.), *The view from building 20*: *Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: The MIT Press.

Halle, Morris, & Alec Marantz. 1994. Some keys of Distributed Morphology. *MIT Working Papers in Linguistics* 21: 275–88.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and Fission. *MIT Working Papers in Linguistics* 30: 425–39.

Harris, James. 1999. Nasal depalatalization no, morphological wellformedness sí: The structure of Spanish word classes. In *Papers on Morphology and Syntax, Cycle One*, 33: 47–82. MIT Working Papers in Linguistics. Cambridge, MA: MIT Press.

Kato, Mary, Martins Ana Maria, Nunes Jairo. 2023. *The Syntax of Portuguese*. Cambridge, UK: Cambridge University Press.

Kelly, Barbara, Gillian Wigglesworth, Rachel Nordlinger, & Joseph Blythe. 2014. The Acquisition of Polysynthetic Languages. *Language and Linguistics Compass* 8(2): 51–64. https://doi.org/10.1111/lnc3.12062

Klima, Edward S. 1964. Relatedness between Grammatical Systems. *Language* 40(1): 1–20. https://doi.org/10.2307/411919

Kučerová, Ivona. 2018. ϕ-features at the syntax-semantics interface: Evidence from nominal inflection. *Linguistic Inquiry* 49(4): 813–845. https://doi.org/10.1162/ling_a_00290

Lees, Robert B. 1957. Syntactic Structures. *Language* 33(3): 375–408. https://doi.org/10.2307/411160

Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge, UK: Cambridge University Press.

Lorandi, Aline. 2010. Formas morfológicas variantes na aquisição da morfologia: evidências da sensibilidade da criança à gramática da língua. *Letrônica* 3(1): 81–96.

MacWhinney, Brian. 2014. *The CHILDES project: Tools for analyzing talk, Volume II: the database (3rd ed.).* New York: Psychology Press. https://doi.org/10.4324/9781315805641

Maldonade, Irani Rodrigues. 2003. *Erros na aquisição da flexão verbal: Uma análise interacionista*. PhD dissertation, Campinas: Universidade Estadual de Campinas.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *U. Penn Working Papers in Linguistics* 4(2): 202–225.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, Fei Xu, & Harald Clahsen. 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57(4): 1–178. https://doi.org/10.2307/1166115

Markman, Ellen M. 1990. Constraints children place on word meanings. *Cognitive Science* 14(1): 57–77. https://doi.org/10.1207/s15516709cog1401_4

Minussi, Rafael Dias; Bassani, Indaiá de Santana. 2017. Em favor do conteúdo semântico das raízes. Revista Letras, Curitiba, UFPR, n. 96, p. 152-173, jul./dez.

Name, Maria. Cristina Lobo. 2008. Pistas prosódicas, sintáticas e semânticas facilitadoras da identificação dos elementos das categorias N e ADJ. In *XXII Encontro Nacional da Anpoll.*

Niyogi, Partha, & Robert C. Berwick. 1995. *The logical problem of language change*. A.I. Memo No. 1516, Cambridge, MA: MIT Artificial Intelligence Laboratory.

Nunes, Rosane. 2003. *Evolução cíclica do futuro do presente do latim ao português*. Master's dissertation, Pelotas: Universidade Católica de Pelotas.

Nunes, Jairo. 2020. Especificação morfológica de pronomes nominativos, concordância verbal e sujeitos nulos em português brasileiro. *Fórum Linguístico* 17(Núm. Esp.): 4658–4672.

Olsen, Susan, Rochelle Lieber, & Pavol Stekauer (eds.). 2014. 'Delineating derivation and compounding'. In *The Oxford handbook of derivational morphology*, 26–49. Oxford: Oxford University Press.

Pearl, Lisa. 2010. Using computational modeling in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research,* Vol. 27: 163–184. Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.27.10pea

Pinker, Steven. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Pinker, Steven, & Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28: 73–193.

Quine, Willard van Norman. 1960. *Word and object*. Cambridge, MA: The MIT Press.

Rainer, Franz. 2005. Semantic change in word formation. *Linguistics* 43(2): 415–441. https://doi.org/10.1515/ling.2005.43.2.415

Resende, Maurício Sartori, & Rodolfo Ilari. 2020. Mudança semântica e formação de palavras. *Letras*, no. 60, 247–54. https://doi.org/10.5902/2176148543108

Resende, Maurício Sartori, & Beatriz Pires Santana. 2019. A relação entre raízes, gênero, classe e significado. *Revista da ABRALIN* 18(1). https://doi.org/10.25189/rabralin.v18i1.481

Resende, Maurício. 2020. Mudança semântica no domínio intravocabular: o caso das raízes cranberry do português. In: Ilari, Rodolfo; Basso, Renato (org.). História do português brasileiro: história semântica do português brasileiro, 120-143. São Paulo: Contexto.

Resende, Maurício Sartori. 2021a. Bebês são não lexicalistas: considerações sobre o estatuto de palavra na aquisição da linguagem. *Diadorim* 23(2): 401–420. https://doi.org/10.35520/diadorim.2021.v23n2a40547

Resende, Maurício Sartori. 2021b. As palavras que os bebês não dizem: revisitando o problema da noção de “palavra” à luz da aquisição da linguagem. *Revista do GEL* 18(2): 128–159. http://dx.doi.org/10.21165/gel.v18i2.3134

Sandmann, Antonio José. 2010. A expressão da pejoratividade. *Revista Letras* 38. https://doi.org/10.5380/rel.v38i0.19181

Santos, Ana Lúcia, & Ruth E. V. Lopes. 2017. Primeiros passos na aquisição da sintaxe: direcionalidade, movimento do verbo e flexão. In A. L. Santos & M. J. Freitas (eds.), *Aquisição de língua materna e não materna: questões gerais e dados do português*, 155–175. Berlin: Language Science Press. https://doi.org/10.5281/ZENODO.889429

Siddiqi, Daniel. 2009. S*yntax within the word: economy, allomorphy, and argument selection in Distributed Morphology*. Linguistics Today 138. Amsterdam; Philadelphia: John Benjamins.

Siskind, Jeffrey Mark. 1996. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition* 61(1–2): 39–91. https://doi.org/10.1016/S0010-0277(96)00728-7

Stevens, Jon Scott, Gleitman, Lila, Trueswell, John C., & Charles Yang. 2017. The pursuit of word meanings. *Cognitive Science* 41: 638–676. https://doi.org/10.1111/cogs.12416

Takahira, Aline Garcia Rodero. 2013. O processo de aquisição de verbos irregulares no português brasileiro. *Estudos Linguísticos* 42(1): 430–441.

Teixeira, Luciana, & Letícia Sicuro Corrêa. 2008. Pistas morfológicas e sintáticas na delimitação de adjetivos em relações predicativas e de adjunção na aquisição do PB. *Revista da ABRALIN*. https://revista.abralin.org/index.php/abralin/article/view/990

Trueswell, John C., Medina, Tamara Nicol, Hafri, Alon, & Lila Gleitman. 2013. Propose but Verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology* 66(1): 126–156. https://doi.org/10.1016/j.cogpsych.2012.10.001

Wexler, Kenneth, & Peter W. Culicover. 1983. *Formal principles of language acquisition*. Cambridge, MA: The MIT Press. https://www.jstor.org/stable/414105

Wuerges, Talita Ewald. 2014. *A aquisição da morfologia verbal por crianças falantes de português brasileiro e o uso de formas variantes*. Masters' thesis. São Paulo: Universidade de São Paulo.

Yang, Charles D. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

Yang, Charles D. 2016. *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: The MIT Press.

Yang, Charles, & Tom Roeper. 2011. *Minimalism and language acquisition*. Manuscript.